



141 aAlaGlnGlyArgAspAspArgProProArgArgGlyGluHisGlyProA 158  
538 AGCACCGGACAACGGCTTGCCGACCGTTTCCCAATGCGGGCGGTATGCT 587  
:||||: ||| ||||| :||  
158 rgHisGlnPro.....ProLeuArgArgArgProGly 168  
588 GAGCAAGAGGAGTAGGCGACGGATTCAAACGCGCCACCCGATACAGCCCG 637  
:||||: :||| :||| :|||: ||  
169 GlyProGlyAlaArgAlaLeuValAlaGluLeuArgProValGlyLeuArg 185  
638 AGCTGGACACATCGGCAATGCCCGCGAAGCCTTCAACGGCAGCTGCAGAT 687  
| ..... |||  
185 g.....ArgV 187  
688 ATCGTCAAAACACATCATCGCGCGGAGGAGAAAT..... 722  
||||: :|||: ||||| :|||: |||  
187 alArgAspAlaArgArgArgArgArgArgAspProArgProTyrPro 203  
723 .....TCTGGCGGAGCGATGCGGTGCAGGTTA 751  
204 ArgLeuGlySerGlyAlaLeuAlaArgAlaArgGlyAlaArgAlaGlyAs 220  
:||||: :||| ||| :|||: |||||  
752 TAAGCGAAGGCTCAACATGCTGCTCATGCACGGCTTGCTGCTTCC 801  
:||| :||| :|||: |||||  
220 pArg.....ValGluLeuGlyProGlyAlaA 229  
802 ACCGAAACAAAGATGGCGCATCAACGATTTGGCAGATATGGCGCACT 851  
||| ||| ||||| |||||: |||  
229 spGlyAspAlaHisGlyArgValProArgArgGlyArgProGlyAlaVal 245  
852 .....CAAGACTA.....TGCGG 865  
246 ValAlaProProArgHisAspGluArgLeuAspProAlaGluAlaPr 262  
866 CAGCAGCATCCGCGATTGGCAGTCCAAAA.....CCCCAATGCCGACACA 912  
262 oArgSerHisProArg...GlyLeuProArgAlaProArgArgGluProA 278  
913 GGCATAGAAGCCGTTCAGCATATCTT..... 938  
|||: |||: |||||  
278 rgArgArgAspArgGlyValAspLeuValAspArgProProAspArgGly 294  
939 .....TATGGCAGCATCCCATCAAGGATTTGGAGCTGCTCC 976  
:||||: ||||| :||| :|||  
295 ArgArgProGlyValAlaGlnHisProLeuArgProProAlaArgGluPr 311  
977 GGGGAAATACGGCTTGCGCGCATCACGSCACATCTGTCAAGCGGTGCG 1026  
|||: :||| ||| ||| ||| :|||: |||  
311 oAlaHisLeuArgAlaArgArgGlyAlaGly.....AlaValAs 324  
1027 CAGATGGCGCGATTCGATTGGCGAAGGAAATCGCGCTCAGCAGCAA 1076  
:||||: ||| :||| :|||: |||||  
324 erAspArgGlyProArg.....ArgAspProHisArgArgAspArgArg 338  
1077 TTTTTCGATGCGGCATACGCCAAATACCGCTCCCTTACCATTCCCGAA 1126  
:||||: ||||| :||| :|||: |||  
339 ArgAlaArgIleLeuAlaArgArgGlyAlaAspProGlyAlaValProGl 355  
1127 ATATCCGTTCAACTTGGACGCGGTTCGCCAAGAAACATCATCTCC 1176  
: ||| ||||| :|||: |||||  
355 uAlaProHisAspArgArgAlaAlaValGlnAspArgArgProGlyProL 372  
1177 TCAACCTGCGCGTCAACACGGCAAAATGTCAAACTGGCAGCAACAACG 1226  
||| ||| ||| :|||: |||||  
372 euLeuArgArgArgAspHisArg...AlaAlaArgAlaHisArgProSer 387  
1227 CCACCCGAAGACAGCGCTACCGTTTGACGGTAAAGGTTTCCGAATTTG 1276  
:||| :||| :|||: |||||  
388 GlyGluAspProArg..LeuProHisArgAlaArg..... 398  
1277 AGAAGCAGGTGAATATGATACGAGCTCGATATTCAGAAATATTTCGGG 1326  
399 .....ArgAsp 400







```
235 snLeuIleArgGluLysGlyLysValThrTrpGlyLeuValTyrAlaThr 251
    :
    :
    :
600 AGCGCAGGATTCAACACGGCCACCGATACAGCCCGAGCTGGACAGAT 649
    ||||| ||| |||:
    :
252 ArgArgLeuSerLeuGlyArgValProValArgLeu.....SerVa 265
    :
    :
650 CGGGCAATCGCCGGAAGCTTCAACGGGCACATGCAGATATCGT..... 692
    :
    :
265 lArgAspGlyArgGlnAlaGlnGluGlnHisSerGluHisArgLeuValG 282
    :
    :
693 .....CAAAAACATCATCGCGCGGCGAGAGAAATTCGCG 728
    :
    :
282 luValGlyGlnGlySerLeuGlnHisLysGluGlyThrArg.....Gln 296
    :
    :
729 CGCAGCGATCGCGTCAGGGTATAAGCA..... 758
    ||||| :|||:
297 ArgArgProAlaArgGlyGlyAspLysGlnLeuArgThrLeuArgGluG 313
    :
    :
759 .....AGGCTCAACAT..... 770
    :
    :
313 ySerProArgGlnArgProArgSerGluArgLeuGlnAspTrpAsnA 330
    :
    :
771 .....TGCTGTCATGCACGCGCT 788
    :
    :
330 rGValGluGlnAspLeuSerLeuAlaAsnValValCysGlyGlyArgAla 346
    :
    :
789 GGGTCTCTTCCACCGA.....AAACAAGATGGCGCGCATCAACGAT 832
    ||||| :|||:
347 GlyGlnLeuArgThrArgGluGlyArgGlnAsnArgHisAlaArgAr 363
    :
    :
833 TGGCAGATATGGCAACTCAAGACTATGCCGAGCAGCATCCGCGAT 882
    :
    :
363 gAlaArgAspSerGluSerGlyAspSerLeuLeuProProArgT 380
    :
    :
883 TGGCAGTCCAAAACCCATCGCGCAACAGCATAGA..... 920
    :
    :
380 yArgAlaProGlnGlyAlaGlyLeuGlnGlyHisArgGluProGlnPro 396
    :
    :
921 .....AGCGTCAGCAATATCTTTATGCAGCCA 949
    :
    :
397 LeuHisAlaProProLeuAlaSerArgSerAsnArgGluGlyGluG 413
    :
    :
950 TCCCATCAAAAGGATTGGAGCTGTCGGGGGAAATACGGCTTGGCGG. 998
    ||||| |||||
413 yProHisGlnArgAspTrp.....LeuGlyArgA 423
    :
    :
999 .....CATCAGGCACATCTCTCAACGGTTCGAGATGGCGGATCGC 1043
    :
    :
423 laGlyGluArgGlyGlyValArgGlnValArgGlyValHisArgGlu 439
    :
    :
1044 ATT...GCCGAAAGGAAATCCGCGT...CAGCGACAATTTTCCCGATG 1087
    :
    :
440 ThrArgGlyProArgTyrValGluHisLeuGlnArgAlaAspGlyArgCy 456
    :
    :
1088 CGGCATACGCCAAATACCGCTCCCTTACCATTCCCGAAATATCCGTTC 1137
    ||||| |||||
456 sGlyAlaArg.....LeuPro.ArgAlaLeuLeuArg... 466
    :
    :
1138 AACTTGGAGCAGCGTTACGGCAAGAAACATACCTCTCAACCGTGCC 1187
    :
    :
467 .....LeuSerAlaGlyGlyTyrG 473
    :
    :
1188 GCGGTCAACAGCGCAAAATGTCAACTGGCAGACCAACGCCCGCGAAGA 1237
    ||||| :|||:
473 uProArgGlyGlyLysAlaGlyAsnProGlnHisAspLysArgProArgT 490
    :
    :
1238 CAGCGGTACCGTTTGCAC.....GGTAAGGGTTT 1266
    :
    :
490 hrGlyLeuGlnAspAspLysGluValArgGlnGlyLysGlyArgGlyPhe 506
    :
    :
1267 CGS.....AATTTTGAGAGACGACGTGAATATGA 1295
    :
    :
507 ProLeuArgGlyArgGlyArgAspAsnLeuGlnHisArgArgCys... 522
1296 TACGAAGCTCGATATT..... 1311
    :
    :
523 .....LeuSerIleArgLeuGlnArgProLysGlyArgGluSerCysA 537
    :
    :
1312 .....CAAGAAATATCGGGGGCGGTATACCTAAGGCTAAGCCTGTG 1353
    ||||| :|||:
537 rGlyArgGlnLeuLeuProGlnArgAlaLeuLeuArgAsnProGln 553
    :
    :
1354 TTTGATCGCAACCGAGATCGGAGTTTCATAGGAAGCTTAATAAATTGAC 1403
    :
    :
554 GlyGlnAlaGlnHisArgValArgAspLeuArgGlnSerSerAlaSe 570
    :
    :
1404 AACTCGTCAGCAGCTGGAGAAAATGTTTCAGGAAACGAGAGAAGAGT. 1452
    :
    :
570 rGlnGlyGluArgLeuAspArgLeuLeuHisLysArgSerArgGlnV 587
    :
    :
1453 .....CAGAGT 1458
    :
    :
587 alPheGlyAlaGlnValProGluHisThrProAspIleLeuProGlySer 603
    :
    :
1459 AGTCAGTTTAAAGCCCATGCCAA...CGAGAATGGGAAAATAAAACAGG 1505
    :
    :
604 SerGlnLeuArgLeuArgLeuGlnAlaArgGluPheArgArgArgLy 620
    :
    :
1506 GTTAGATTTTAAATCATTTTATAGTGTGTATCAATAAGAAAGGCACAG 1555
    :
    :
620 salaArgLysArgHisArgLeuGly..... 628
    :
    :
1556 TAACAGGAGGCGATAGTACCCGCTGTGTGTACGGGTGATACACAACAA 1605
    :
    :
629 .....AspLeuSerGlyGlyAspLeuArgLeu..... 637
    :
    :
1606 ACCTCGCACCTGATAAAACATGGGCTTTTATCAAGCGACAGTGGAAATTA 1655
    :
    :
638 .....AspLysArgGly.....GlnGlnIleAr 645
    :
    :
1656 AAAGCGCTGATGGAATGTGGAGGTGAAACGAAAAA.....GGTG 1696
    :
    :
645 gGlyProGlyLeuArgHisArgLysArgAsnSerArgPheAsnHisProA 662
    :
    :
1697 GGAAGATGATGACCAAGCAGCACCATGTTCCCAAAAGATTGGGATGAGCT 1746
    :
    :
662 laAlaValLeu.ProArgGluProCysSerGluAspGlyGlyValArgG 678
    :
    :
1747 AGAATTA.....GGCTGAAGTTACTT 1768
    :
    :
678 yGlyLeuArgArgGlnGlyLeuProLeuLeu 688

seq_name: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:US-08-760-489-4

seq_documentation_block:
; Sequence 4, Application US/08760489
; Patent No. 5830696
; GENERAL INFORMATION:
; APPLICANT: Short, Jay M.
; TITLE OF INVENTION: DIRECTED EVOLUTION OF THERMOPHILIC
; TITLE OF INVENTION: ENZYMES
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
```











```
|||||
252 ArgArgLeuSerLeuGlyArgValProValArgLeu.....SerVa 265
650 CGGGCAATGCCCGGAGCTTCAACGGCAGTGCAGATATCGT..... 692
265 IArgaspGlyArgGlnAlaGlnGluGluHisSerGluHisArgLeuValG 282
693 .....CAAAACATCATCGCGCGGAGGAGAAATGTGCGG 728
282 luValGlyGlnGlySerLeuGlnHisLysGluGlyThrArg.....Gln 296
729 CGCAGCGCATCGCGTGCAGGGTATAAGCGA..... 758
297 ArgArgProAlaArgGlyGlyAspLysGlnLeuArgThrLeuArgGluG 313
759 .....AGGCTCAACAT..... 770
313 ySerProProArgGlnArgProArgSerGluArgLeuGlnSerTrpAsnA 330
771 .....TGCTGTATCGACGGCTT 788
330 rgValGluGlnAspLeuSerLeuAlaAsnValValCysGlyGlyArgAla 346
789 GGTCTGCTTCCACCGA.....AAACAAGATGCGCGCATCAACGATT 832
347 GlyGlnLeuArgThrArgGluGlyArgGlnAsnArgArgHisAlaArgA 363
833 TGGCAGATATGCGCACTCAAGACTATGCGCGCAGCAGCATCGCGCAT 882
363 gAlaArgArgAspSerGluSerGlyAspSerLeuLeuProProArgT 380
883 TGGCAGTCCAAACCCCAATCCGCACAAAGGCATAGA..... 920
380 yrArgAlaProGlnGlyAlaGlyLeuGlnGlyHisArgGluProGlnPro 396
921 .....AGCGTCAAGCAATATCTTTATGCGAGCCA 949
397 LeuHisAlaProProLeuAlaSerArgSerAspAsnArgGluGlyGluG 413
950 TCCCATCAAGGGATTGGAGTGTCCGGGAAATACGGCTTGGCGG. 998
413 yProHisGlnArgAspTrp.....LeuGlyArgA 423
999 .....CATCAGCGCATCTCTCAAGCGTGCAGATGCGCGCATCGC 1043
423 laGlyGluArgGlyGlyValArgGlnValArgGlyValHisArgGluArg 439
1044 ATT...GCCGAAGGAATCCGCGT...CAGCGACAATTTGCCGATG 1087
440 ThrArgGlyProArgTyrValGluHisLeuGlnArgAlaAspGlyArgCy 456
1088 CGGCATACGCCAAATACCGTCCCTTACCATTCGCCAAATATCGTTCA 1137
456 sGlyAlaArg.....LeuPro.ArgAlaLeuLeuArg... 466
1138 AACTGGAGCAGCTTACGGCAAGAAACATCACTCTCAACCGTGCC 1187
467 .....LeuSerAlaGlyGlyTyrG 473
1188 GCGTCAACGCCAAATATGCTCAACTGCAGACCAACGCCACCGCAAGA 1237
473 uProArgGlyGlyLysAlaGlyAsnProGlnHisAspLysArgProArgT 490
1238 CAGGGTACCGTTTGAC.....GGTAAGGGTTT 1266
490 hrGlyLeuGlnAspAspLysGluValArgGlnGlyLysGlyArgGlyphe 506
1267 CCG.....AATTGAGAGCAGCAGTCAATATGA 1295
507 ProLeuArgGlyArgGlyArgAspAsnLeuGlnGlnHisArgArgCys... 522
1296 TAGAAGCTCGATATT..... 1311
|||||

seq_name: /cqn2.6/ptodata/1/1aa/6B_COMB.pap:US-09-199-637A-351
seq_documentation_block:
; Sequence 351, Application US/09199637A
; Patent No. 6355411
; GENERAL INFORMATION:
; APPLICANT: Ausubel, Frederick
; APPLICANT: Goodman, Howard M.
; APPLICANT: Rahme, Laurence G.
; APPLICANT: Mahajan-Miklos, Shalina
; APPLICANT: Tan, Man-Wah
; APPLICANT: Cao, Hui
; APPLICANT: Drenkard, Eliana
; APPLICANT: Tsongalis, John
; TITLE OF INVENTION: VIRULENCE-ASSOCIATED NUCLEIC ACID
; TITLE OF INVENTION: SEQUENCES AND USES THEREOF
; FILE REFERENCE: 00786/361002
; CURRENT FILING DATE: US/09/199,637A
; PRIOR APPLICATION NUMBER: 60/066,517
; PRIOR FILING DATE: 1997-11-25
; NUMBER OF SEQ ID NOS: 437
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 351
; LENGTH: 447
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
```

US-09-199-637A-351

alignment\_scores:  
Quality: 125.50 Length: 442  
Ratio: 0.743 Gaps: 20  
Percent Similarity: 38.235 Percent Identity: 23.982

alignment\_block:

US-09-303-518D-463 x US-09-199-637A-351 ..

Align seg 1/1 to: US-09-199-637A-351 from: 1 to: 447

```
99 CCGCGAGGTTCTCGACCGTCAGCATTTCCGACCCGACGGGAATACCACC 148
|||||
7 ProAlaThrSerArgLysProArgValAlaGlyProArgAlaValArgPr 23
149 TATTCGGCAGCGGGGAGCTGTCNACCGCAACGGCCATATCGGATG 198
|||||
23 OValProHisArgHisAlaProAlaGlyArgGlnArgProArgSerGlyP 40
199 GGAACAT...ACAAAGCCATCAGTTGGCCACCTGATGATTCACAGGC 245
|||||
40 roValHisProAlaAlaProSerThrAspPro.....Gly 51
246 GCGCGTTGAAGGAATATCGGCTACATTTCCGCTTTCCGATCACGGGC 295
|||||
52 Gly.....HisAspHisArgValArgGlnProGlyHisArgAspProG 66
296 ACNAATCCATTCGCCCTTCGACACCATTCGCTCACAATTCGATTCGAC 345
|||||
66 yAlaGlnGlyArgCysLeuArg..... 73
346 GAAGCCGGTAGTCCCGTTGACGGATTACCGCTTACCGCATCCATTTGGA 395
|||||
74 .....LeuProHisGlnThrGly 79
396 CGGATACGAACACC.....ATCCCGCCGACGGCTAT 426
|||||
80 ArgProArgSerLeuAlaGlyAlaGlyGlyAsnArgProThrLeuAlaG 96
427 GACGGCCACAGCGCGGCTATCCCGTCCCAAGCGCGGAGGATAT 476
|||||
96 nProGlySerArgGlySerAlaGlyGlnProProAlaArgArgVal. 112
477 ATACAGCTACGACATAAAGCGTGTGCCCAAAATATCCGCTCAACCTGA 526
|||||
113 .....AlaAlaAspAlaArgProAlaGlnProAs 122
527 CGGACACCGCAGCACCGGACACCGCTTGGCCGCTTTCCACATGCC 576
|||||
122 pArgGln.....AlaGlyAlaGlnProGlyAlaGlyLeuH 134
577 GCGCTATGCTGACGCAAGAGTAGCGGACGATTCACACGCGCCACCG 626
|||||
134 isGlnTrp.....ArgValArgGlnArgGlnGlyThrGlyGlyAla 147
627 ATACAGCCCGAGCTGGA.....CAGATCGGCA 655
|||||
148 ProAspProArgAlaGlyAlaThrTyArgAlaAlaValArgAlaGlyG 164
656 ATCCGCGCGA..... 665
164 uLeuArgArgAspSerLeuArgAlaAspGlyLysArgValLeuArgPro 181
666 ..AGCCTTCAACGGCACTCAGATAT..... 689
181 lngluArgGlnLeuHisTrpArgTyArgGlnAlaGlyProValPro 197
690 ...CGTCAAAACATCAT.....CGCGCGCAGGAGAAATGT... 725
|||||
198 GlyArgGlnArgTrpHisProValProArgArgSerArgProAlaAs 214
```

```
726 .....CGGC 729
214 pGlyHisAlaGlyGlnThrAlaProGlyAspProGlyLysGlyAlaA 231
730 GCAGCGCATGCCGTGACGGTATAGCGAAGGCTCAAAACATTGCTGTCAT 779
|||||
231 rgGlyArgArgProAlaGlyGlyArgArgArgThrCysAlaSerSerAla 247
780 GCACGGCTTGGTCTGCTTCCACCCGAAACAAAGATGGCGCGCATCAACG 829
|||||
248 ProProThrArgThrSerProProLysSerAlaProGlyAlaSerAlaA 264
830 ATTTGGCAGATATGGCGCAACTCAAGACTATGCCGCGCAGCGCATCCGC 879
|||||
264 gThrSerThrThrAlaSerThrSerSerCysAlaTyThrAlaAlaA 281
880 GATTGGCAGTCCAAACCCCAATGCCGACAAAGGCATAGAACCGCTCAG 929
|||||
281 rgThrProArgGlyHisProAlaAlaArg..... 291
930 CAATATCTTTATGGCAGCATCCCATCAAGGATTTGGAGCTGTCGGG 979
|||||
292 .....ThrHisPro.....GlnAlaProG 298
980 GAAATACGCTTGGCGGCGCATCACGCACATCTCTCAAGCGGTCCGAG 1029
|||||
298 yArg.....ArgHisArgProAlaGlyArgGlnAla..... 308
1030 ATGGCGCGATCGCATTCGCGAAAGGAAATCCGCGCTCAGCGACAATTT 1079
|||||
309 ..AspArgArgArgThrGlyGlnAlaGlu..... 317
1080 TGGCGATCGCGCATACGCCAAATACCGTCCCTTACCATTTCCGAAATA 1129
|||||
318 .....GluLeuProLeuProGlyGlnArgProArgAl 328
1130 TCCGTTCAAACTTGGACGAGCTTA.....CGCAAGAAACATC 1170
|||||
328 agLysHisAlaGlyAlaArgLeuTyProValArgArgProAspp 345
1171 ACCTCTCAACCGTCCGCGCTCAACCGCAAAATGTCAAACCTGGCAGA 1220
|||||
345 roAlaSerArgProAlaProGlyArgCysAlaGlyCysGlnProGlyArg 361
1221 CCAACGCGCACCGCAAGACAGCGCT 1244
362 ArgArgGluProGluArgAsnArg 369
seq_name: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:US-09-007-999-2
seq_documentation_block:
; Sequence 2, Application US/09007999
; Patent No. 6087559
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starch and
; FILE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0356D
; CURRENT APPLICATION NUMBER: US/09/007,999
; CURRENT FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
US-09-007-999-2
alignment_scores: Quality: 125.00 Length: 751
```

Ratio: 0.401 Gaps: 46  
Percent Similarity: 41.545 Percent Identity: 21.571  
alignment\_block:  
US-09-303-518D-463 x US-09-007-999-2 ..  
Align seg 1/1 to: US-09-007-999-2 from: 1 to: 1475

64 GCACAGCCCTCAGATTGGCAACAGCATCCCTTTATCCGGCAGGTTCTCGA 113  
:::||||| |||:::|||||  
617 SerTyrAlaLeuLeuLeuThrasnLysSerValProArgValTyrTy 633  
:::||||| |||:::|||||  
114 CGCTCAGCATTCGCAACCCGCGGAAATAC..... 144  
::: ||| |||||:::|||||  
633 rGlyAspMetPheThrAspGlyGlnTyrMetAlaHisLysThrIleA 650  
145 .....CACCTATTCCGCAGCAGG..... 162  
650 snTyrGluAlaIleGluThrLeuLeuLysAlaArgIleLysTyrValSer 666  
163 GGGAGCTTGGCCAGCGCAACGCCATATC..... 192  
||| ||| |||||:::|||||  
667 GlyGlyGlnAlaMetArgAsnGlnGlnValGlyAsnSerGluIleIleTh 683  
193 .....GGATTGGGAAACATACAAAGCCATCAGTTGGGCCACC 229  
||| ||| |||:::|||||  
683 rSerValArgTyrGlyLysGlyAlaLeuLysAlaThrAspThrGlyAspA 700  
230 TGATGATTCACACAG.....GGCGCGTTGAAGGAAATATCGGCTAC 270  
700 rgThrThrArgThrSerGlyValAlaValIleGluGlyAsnAsnProSer 716  
271 ATT...GTCCGCTTTCCGAT.....CA 290  
::: ||| |||||  
717 LeuArgLeuLysAlaSerAspArgValValValAsnMetGlyAlaAlaHi 733  
291 CGGGCACAAATCCATTCGCCCTTC.....GACAAAC.....C 322  
||| ||| |||:::|||||  
733 sLysAsnGlnAlaTyrArgProLeuLeuLeuThrThrAspAsnGlyIleL 750  
323 ATGCCTCATCTCCGATTCCTGACGAAGCCGGTAGTCCCGTTGACGGATTTC 372  
||| ||| |||||:::|||||  
750 ysAlaTyrHisSerAspGlnGluAlaGly..... 760  
373 ACCCTTTACCGCATCCATTGGAGGATACGAA.....CACCATCCCGC 416  
||| ||| |||  
761 ...LeuValArgTyrThrAsnAspArgGlyGluLeuIlePheThrAlaAl 776  
417 CGAC.....GGCTATGACGGCCACAGCGCGGCTATCCC..... 453  
||| ||| |||||:::|||||  
776 asPileLysGlyTyrAlaAsnProGlnValSerGlyTyrLeuGlyValT 793  
454 ..GTCCCAAAAGCGCGAGGATATATACAGCTACGACATAAAAGCGTT 501  
:::||||| ||| |||:::|||||  
793 rpValProValGlyAlaAlaLeuIleLysMetPheAlaLeuArgLeuAla 809  
502 GCCCAAAATATCCGCTCAACCTGACCGCAACCGCAGCAGCCGACACG 551  
::: ||| |||||  
810 ArgProHis.....GlnG 814  
552 GCTTGGCCAGCGTTTCCACAAATGCCGGCTATCTCTGACGCAAGGATAG 601  
:::||||| ||| |||:::|||||  
814 nMetAlaSerValHisGlnAsnAla...AlaLeuAspSerArgValMetP 830  
602 CGCAGGATTCAAACCGCCACCGCATACAGCCCGGAGCTGGACAGATCG 651  
:::||||| ||| |||:::|||||  
830 heGluGlyPheSerAsnPheGlnAlaPheAlaThrLysLysGluGluTyr 846  
652 GGCAATGCCCGCGAGCTTCAAC.....GGCAC 680  
|||:::||||| ||| |||  
847 ThrAsnValValIleAlaLysAsnValAspLysPheAlaGluTrpGlyVa 863  
681 TGCAGATATCTCAAAAACATCATCTCGCGCGGAGGAGAAATGTCGCGC 730

863 lThrAsp.....PheGluMetAlaProGlnTyrValSerS 875  
:::||||| |||:::|||||  
731 CAGCGCATGCCGTG.....CAGGCTATAAGCAAGGCTCAACATT 771  
::: |||||:::|||||  
875 erThrAspGlySerPheLeuAspSerValIleGlnAsnGlyTyrAlaPhe 891  
772 GCTGTCTATGCACGCTTGGTCTCTCTTCCACCGAAACAAAGATGCGCG 821  
::: |||||:::|||||  
892 ThrAspArgTyrAspLeuGly...IleSerLysProAsnLysTyrGlyTh 907  
822 CATCAACGATTTGGCAGAT...ATGGCGCAACTCAAAAGACTATCCGCCAG 868  
:::|||||:::|||||  
907 rAlaAspAspLeuValLysAlaIleLysAlaLeuHisserLysGlyIleL 924  
869 CAGCCATCCCGGATTGGCGAGTCCAAACCCCAATGCCGCACAGGCATA 918  
::: |||||:::|||||  
924 ysValMetAlaAspTrpValProAspGln..... 933  
919 GAAGCCCTCAGCAATATCTTTATGGCAGCCATCCCATCAAGGG...AT 965  
||| |||:::|||||  
934 .....MetTyrAlaPheProGluLysGluValVa 943  
966 TGGAGCTGTCCGG...GGAATAACGGCTTGGCGGCATCACGCCACATC 1012  
::: |||||:::|||||  
943 iThrAlaThrArgValAspLysTyrGlyThr.....P 954  
1013 CTGTCAACGCGTCCGAGATG...GGCGCATCGCATTTGCCGAAAGGAAA 1059  
||| |||||:::|||||  
954 roValAlaGlySerGlnIleLysAsnThrLeuTyrValValAspGlyLys 970  
1060 TCGCGCTCAGCGACAT.....TTTCCCGATGC 1088  
|||::: |||||:::|||||  
971 SerSerGlyLysAspGlnGlnAlaLysTyrGlyAlaPheLeuGlu 987  
1089 GGCATACGCCAAATACCCCTCTTACCATTCCCGAAATATCCGTCAA 1138  
||| |||||:::|||||  
987 uLeuGlnAlaLysTyrProGluLeuPheAlaArgLysGlnIle..... 1001  
1139 ACTTGGAGCAGCTTACGGCAAGAAACATCACCTCTCAACCCGTCGCG 1188  
|||::: |||||  
1002 .....SerThrGlyValPro 1006  
1189 .....CCGTCA.....AACGG 1199  
||| |||||  
1007 MetAspProSerValLysIleLysGlnTrpSerAlaLysTyrPheAsnG 1023  
1200 CAAAAATGTC..... 1209  
|||::: |||||  
1023 yThrAsnIleLeuGlyArgGlyAlaGlyTyrValLeuLysAspGlnAlat 1040  
1210 .....AACTGGCAGACCAAGCCAC.....CCG 1233  
|||::: |||||:::|||||  
1040 hrAsnThrTyrPheAsnIleSerAspAsnLysGluIleAsnPheLeuPro 1056  
1234 AAGACA.....GGCGTACCCTTTTACGCTAA 1259  
||| |||||  
1057 LysThrLeuLeuAsnGlnAspSerGlnValGlyPheSerTyrAspGlyLy 1073  
1260 AGGTTTCCGAATTTTGAGACACGCTGAATATATGATACGAAGCTCGATA 1309  
|||::: |||||  
1073 sGlyTyr.....ValTyrTyrSerThr..... 1080  
1310 TTCAAGAATATCGGGCGCGGTATACCTAAGCTAAGCTGTGTGTGT 1359  
||| |||||  
1081 .....SerGly.....TyrGlnAlaLysAsnThrPheIle 1090  
1360 GCGAAA...CCGAGATGG...GAGGTTGATAGGAAG..... 1389  
|||::: |||||  
1091 SerGluGlyAspLysTrpTyrPheAspAsnAsnGlyTyrMetValTh 1107  
1390 .....CTTAATAAT 1399  
||| |||

```

1107 rGlyAlaGlnSerIleAsnGlyValAsnTyrTyrPheLeuSerAsnGlyL 1124
1400 TGACAACTCGTGAAGAGTGGGAGAAATGTCAGGAACAGAGAGAGG 1449
1124 euGlnLeuArgAspAlaIleLeuLysAsnGluAspGlyThr 1137
1450 AGTCAGAGTAGTCAGTTTAAAGCCCATCGGCAACGAGAGATGGGAAAT 1497
1138 .....TyrAlaTyrTyrGlyAsnAspGlyArgArgTyrGluAsnG 1151
1498 .....AAACAGGGTTAGATTTTAAATCATTTTATAGGTGGTG 1534
1151 yTyrTyrGlnPheMetSerGlyVal...TirArgHisPheAsnAsnGly 1167
1535 ATATCAATAGAAGGACAGTAAACAGGAGGAGGCATAGTCTAACCCGTGT 1584
1167 luMetSerValGlyLeuThrValIleAspGly 1177
1585 GATGTACGGGTGATACAAACAACTCGGCACCTGATATAACATGGGTTTA 1634
1178 .....GlnValGlnTyrPheAspGluMetGly...Ty 1187
1635 TCAAGGCACAGTGGAA...ATTAAAGCCTGATGGAAGTTGGAGGTGA 1681
1187 rGlnAlaLysGlyLysPheValThrAlaAspGlyLysIleArgTyrP 1204
1682 AAACGAAAAGGTGGGAAAGTGTACCAACGACACACCATGTTCCCAAAA 1731
1204 heAspLysGlnSerGlyAsnMet 1211
1732 GATTGGGATGAGGCTAGATAGGCTGAAGTTACTTCGGCTTGGGAAAG 1781
1212 .....TyrArgAs 1214
1782 TAGA...ATATGCTTAAGATATATAATGCGAGGTACAGTAATACG 1828
1214 nArgPheIleGluAsnGluGlyLysTrpLeuTyrLeuGlyGluAspG 1231
1829 GTATTAAATAGAAGGATTTACCGAACCTAATAGAACAGCATATCCCAT 1878
1231 lyAlaAlaValThrGlySerGlnThrIleAsn.GlyGlnHisLeuTyrPh 1247
1879 T 1879
1247 e 1247

```

seq\_name: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:us-09-210-361-2

```

seq_documentation_block:
; Sequence 2, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2

```

```

; LENGTH: 1475
; TYPE: PRT
; ORGANISM: Streptococcus mutans
; US-09-210-361-2

alignment_scores:
  Quality: 125.00      Length: 751
  Ratio: 0.401        Gaps: 46
  Percent Similarity: 41.545  Percent Identity: 21.571

alignment_block:
US-09-303-518D-463 x US-09-210-361-2  ..

Align seg 1/1 to: US-09-210-361-2 from: 1 to: 1475

64 GCACACGCTCAGATTGGCAACAGATCCCTTTATCCGGCAGGTTCTCGA 113
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
617 SerTyrAlaLeuLeuLeuThrAsnLysSerSerValProArgValTyrTy 633
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
114 CCGTCAGCATTTTCGAACCCGACGGGAATATAC..... 144
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
633 rGlyAspMetPheThrAspAspGlyGlnTyrMetAlaHisLysThrIleA 650
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
145 .....CACCTATTCGCACGACG..... 162
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
650 snTyrGluAlaIleGluThrLeuLeuLysAlaArgIleLysTyrValSer 666
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
163 GGGGAGCTTGCNAGCGCAACGGCCATATC..... 192
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
667 GlyGlyGlnAlaMetArgAsnGlnValGlyAsnSerGluIleIleTh 683
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
193 .....GGATTGGGAACATACAAAGCCATCAGTTGGGCCACC 229
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
683 rSerValArgTyrGlyLysGlyAlaLeuLysAlaThrAspThrGlyAspA 700
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
230 TCATGATTCACACAG.....GGGCGCGTTGAAGGAAATATCGGCTAC 270
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
700 rGThrThrArgThrSerGlyValAlaValIleGluGlyAsnAsnProSer 716
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
271 ATT...GTCCGCTTTTCCGAT.....CA 290
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
717 LeuArgLeuLysAlaSerArgValValValAsnMetGlyAlaAlaHi 733
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
291 CGGCAACAATTCATTCGCCCTTC.....CACAC.....C 322
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
733 sLysAsnGlnAlaTyrArgProLeuLeuLeuThrThrAspAsnGlyIleL 750
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
323 ATGCCTCACATTCCGATTCTGACGAAGCGGTAGTCCCGTTGACGGATTC 372
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
750 ysAlaTyrHisSerAspGlnGluAlaAlaGly..... 760
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
373 AGCCTTTACCGCATCCATTGGGACGGATACGAA.....CACATCCCGC 416
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
761 ...LeuValArgTyrThrAsnAspArgGlyGluLeuIlePheThrAlaAl 776
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
417 CGAC.....GCTATGACGGCCACAGCGCGGCTATCCC..... 453
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
776 aAspIleLysGlyTyrAlaAsnProGlnValSerGlyTyrLeuGlyValT 793
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
454 ..GCTCCCAAGGCGCGAGGGATATATACAGCTACGATACATAAAGCGTT 501
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
793 rpValProValGlyAlaAlaLeuIleLysMetPheAlaLeuArgLeuAla 809
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
502 GCCCAAAATATCCGCTCAACCTGACCCGACACCGCAGCAGCGGACACG 551
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
810 ArgProHis.....GlnG 814
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
552 GCTTCCCGAGCGTTTCCCAAAATCCCGCGCTATGCTGACGCAAGGAGTAG 601
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
814 nMetAlaSerValHisGlnAsnAla...AlaLeuAspSerArgValMetP 830
   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
602 GCGACGGATTCAAACGCGCACCGCCGATACAGCCCGAGCTGGACAGATCG 651

```

1081 .....SerGly.....TyrGlnAlaLysAsnThrPheIle 1090  
1360 GCAGAA...CCGAGATGG...GAGGTTTCATAGGAG 1369  
::: ::::: ::::: |||::: ::::: 1378  
1091 SerGluGlyAspLysTrpTyrTyrPheAspAsnAsnGlyTyrMetValTh 1107  
1390 .....:::.....CTTAATAAAT 1399  
::: ::::: ::::: ||| 1408  
1107 rGlyAlaGlnSerIleAsnGlyValAsnTyrTyrPheLeuSerAsnGlyL 1124  
1400 TGACAACTCGTGACGAGGTGGAGAAAATGTCACGAAACACGAGAAGAAG 1449  
|||::: ::::: |||::: ::::: ||| 1458  
1124 euGlnLeuArgAspAlaIleLeuLysAsnGluAspGlyThr..... 1137  
1450 AGTCAGAGTAGTCAGCTTAAAGCCCATCGCAACAGAAATGGGAAAAT... 1497  
::: ::::: ::::: |||::: |||::: ||| 1506  
1138 .....TyrAlaTyrTyrGlyAsnAspGlyArgArgTyrGluAsnGl 1151  
1498 .....AAACAGGGTTAGATTTTAATTCATTATTATAGGTGGTG 1534  
::: |||::: ::::: |||::: |||::: ||| 1543  
1151 yTyrTyrGlnPheMetSerGlyVal...TrpArgHisPheAsnAsnGlyG 1167  
1535 ATATCAATAAGAAGGACACAGTAAACAGAGGCGCATAGCTAACCCCGTGT 1584  
|||::: |||::: ||| 1593  
1167 luMetSerValCylLeuThrValIleAspGly..... 1177  
1585 GATGTACGGGTGATACAAACAACCTCGGCACCTGATAAATCGGGTTTA 1634  
|||::: |||::: |||::: ||| 1643  
1178 .....GlnValGlnTyrPheAspGluMetGly...Ty 1187  
1635 TCAACGCGACAGTGAA...ATTAAAGACCTGATGGAAAGTTGGGAGGTGA 1681  
|||::: ::::: |||::: |||::: ||| 1690  
1187 rGlnAlaLysGlyLysPheValThrAlaAspGlyLysIleArgTyrP 1204  
1682 AAACCAAAAAAGTGGGAAGTGATGACCAAGCACACCATGTTCCCAAAA 1731  
|||::: |||::: ||| 1740  
1204 heAspLysGlnSerGlyAsnMet..... 1211  
1732 GATTGGGATGAGGCTAGATAATTAGGGCTGAAGTTACTTCGGCTGGGAAAG 1781  
::: ::::: ::::: TyArgAs 1214  
1212 .....:::..... 1211  
1782 TAGA...ATAATGCTTAAGGATATAAATGGCAGGTGACAAGTAAATCGG 1828  
::: ||| ::::: |||::: ||| 1837  
1214 nArgPheIleGluAsnGluGlyLysTrpLeuTyrLeuGlyGluAspG 1231  
1829 GTATTAAATAGAAGGATTTTACCGAACCTAATAGACACCATATCCATT 1878  
|||::: |||::: ||| 1887  
1231 lyAlaAlaValThrGlySerGlnThrIleAsn...GlyGlnHisLeuTyrPh 1247  
1879 T 1879  
I 1247 e 1247

seq\_documentation\_block:

; Sequence 2, Application US/08793824  
: Patent No. 5081939

APPLICANT: Simpson, Christine Lynn

Genetic Manipulation of Plants to

CORRESPONDENCE ADDRESS:

CITY: No. 5981838th Sydney

ZIP: 2060





```
1065 CQTCAGCAGCAATTTTCCGATCGCGCATAGCCAAATACCCGTCCTT 1114
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1092 .....PheLeuAspGluLeuLysAlaLysTyrProAlaIleP 1104
1115 ACCATTCCGGAATATCCGTTCAACTTGGAGCAGCGTTACGGCAAGAA 1164
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1104 heGluArgValGlnIle...SerAsnGlyArgLysLeuThrThrAsnGlu 1119
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1165 AACATCACCTCCTCAACCGCCCGCGTCAACCGCAAAATGCAAACT 1214
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1120 LysIleThrGlnTrpSerAlaLysTyrPheAsnGlySerAsnIleGlnG 1136
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1215 GCGACACCAACGCCAC..... 1230
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1136 yThrGlyAlaArgTyrValLeuGlnAspAsnAlaThrAsnGlnTyrPheS 1153
1231 .....CCGAACACAGCGGTACCGTTTCAC 1254
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1153 erValLysAlaGlyGlnThrPheLeuProLysGlnMetThrGluIleThr 1169
1255 GGTAAAGGTTTCCGAATTTTGAGAAGCACGTGAATATGATACGAAGCT 1304
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1170 GlySerGlyPheArgArgValGlyAsp..... 1178
1305 CGATATTCAAGAATATATCG...GGGGGGGTATACCTAAGGCTAAGCCTG 1351
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1179 AspValGlnTyrLeuSerIleGlyGlyTyrLeuAlaLysAsnThrPheI 1195
1352 TGTTTGATCGGMAACCGAGTGGAGGTGTATAGGACCTTAATAATTG 1401
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1195 leGlnValGlyAlaAsnGlnTrpTyrTyrPheAspLysAsnGlyAsnMet 1211
1402 ACAACTCGTGAGCAGGTG..... 1419
1212 ValThrGlyGluGlnValIleAspGlyLysLysTyrPhePheLeuAspAs 1228
1419 ..... 1419
1228 nGlyLeuGlnLeuArgHisValLeuArgGlnGlySerAspGlyHisValT 1245
1420 .....GAGAAATGTTCCAGAA..... 1437
1245 yTyrTyrAspProLysGlyValGlnAlaPheAsnGlyPheTyrAspPhe 1261
1438 AGCAGAAGAGGAGTCACAGTAGTCAGTTTAAAGCCCATCGCACGA... 1485
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1262 AlaGlyProArgGlnAspValArgTyrPheAspGlyAsnGlyGlnMetTy 1278
1486 .....GAATGGGAAATAAAA 1501
1278 rArgGlyLeuHisAspMetTyrGlyThrThrPheTyrPheAspGluLysT 1295
1502 CAGGG..... 1506
1295 hrGlyIleGlnAlaLysAspLysPheIleArgPheAlaAspGlyArgThr 1311
1507 .....TTAGATTTTATCATTTATAGG 1529
1312 ArgTyrPheIleProAspThrGlyAsnLeuAlaValAsnArgPheAlaG 1328
1530 TGGTGATATCAATAAG..... 1545
1328 nAsnProGluAsnLysAlaTrpTyrTyrLeuAspSerAsnGlyTyrAlav 1345
1546 .....AAA 1548
1345 alThrGlyLeuGlnThrIleAsnGlyLysGlnTyrTyrPheAspAsnGlu 1361
1549 GGCACAGTAACAGGAGGAGTAGCTAACCCTGCTGATGACGGTGAT 1598
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1362 GlyArgGlnValLysGlyHisPheValThrIleAsnAsnGlnArgTyrPh 1378
1599 ACAACAAACCTCG.....GCACCTGATAAACATGGGTTTATC 1636
```

```
1378 eLeuAspGlyAspSerGlyGluIleAlaProSerArgPhe..... 1391
1637 AACGCGACGTGGAATTTAAAAAG.....CCTGATGGAAGTTGGGAG 1677
1392 .....ValThrGluAsnAsnLysTrpTyrTyrValAspGlyAsnGly... 1405
1678 GTGAAACGAAAGAGTGGGAAAGTGTATGACC...AAGCACACCATGTT 1724
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1406 ...LysLeuValLysGlyAlaGlnValIleAsnGlyAsnHisTyrTyrPh 1421
1725 CCCAAAGATGGGATGAGGTAGGCTAGGCTAGGCTAGGCTTACTTCGCTT 1774
      : : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
1421 eAsnAsnAspTyr.....SerGlnValLysGlyValat 1432
1775 GCGAAAGTAGA..... 1785
1432 rPalAsnGlyArgTyrTyrAspGlyAspSerGlyGlnAlaValSerAsn 1448
1786 .....ATAATGCTTAAGGATATAATGCGAGGTACAGGTAAATCGGG 1829
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1449 GlnPheIleGlnIleAlaAlaAsnGlnTrpAlaTyrLeuAsnGlnAspG 1465
1830 TATTAATAAGAGGATTTACCGAACCTATATAGAACACGATAT 1872
1465 yHisLysValThrGlyLeuGlnAsnIleAsnAsnLysValTyr 1479
seq_name: /cgn2_6/ptodata/1/laa/6B_COMB.pep:US-09-210-361-4
seq_documentation_block:
; Sequence 4, Application US/09210361
; Patent No. 6284479
; GENERAL INFORMATION:
; APPLICANT: Nichols, Scott E.
; TITLE OF INVENTION: Substitutes for Modified Starches and
; FILE OF INVENTION: Latexes in Paper Manufacture
; FILE REFERENCE: 0357CR
; CURRENT APPLICATION NUMBER: US/09/210,361
; CURRENT FILING DATE: 1998-12-11
; EARLIER APPLICATION NUMBER: 09/007,999
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/478,704
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/009,620
; EARLIER FILING DATE: 1998-01-20
; EARLIER APPLICATION NUMBER: 08/485,243
; EARLIER FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 09/008,172
; EARLIER FILING DATE: 1998-01-16
; EARLIER APPLICATION NUMBER: 08/482,711
; EARLIER FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4
; LENGTH: 1375
; TYPE: PRT
; ORGANISM: streptococcus mutans
US-09-210-361-4
```

```
alignment_scores:
  Quality: 123.00      Length: 533
  Ratio: 0.557        Gaps: 32
  Percent Similarity: 41.463      Percent Identity: 22.889

alignment_block:
US-09-303-518D-463 x US-09-210-361-4 ..
Align seg 1/1 to: US-09-210-361-4 from: 1 to: 1375
64 GCACACGCTCAGATTGGCAACAGATCCCTTTATCCGGCAGGTCTCGA 113
: : : : : ||| ||| ||| ||| ||| ||| ||| ||| |||
643 SerTyrAlaLeuLeuThrAsnLysSerSerValProArgValTyrTy 659
```

```
114 CCCTCAGCATTTGCAACCCGACCGGAAATAC.....144
    :|||:|||||:|||||:
659 rGlyAspMetPheThrAspAspGlyGlnTyrMetAlaHisLysThrIleA 676
145 .....CACCTATTCCGACGAGG.....162
676 snTyrGluAlaIleGluThrLeuLeuLysAlaAlaArgIleLysTyrValSer 692
163 GGGGAGCTTGCNAGCGCAACGCGCCATATC.....192
693 GlyGlyGlnAlaMetArgAsnGlnGlnValGlyAsnSerGluIleIleTh 709
193 .....GGATTGGGAACATACAAAGCCATCAGTTGGGCGAC 229
709 rSerValArgTyrGlyLysGlyAlaLeuLysAlaThrAspThrGlyAspA 726
230 TGATGATTCACAG.....CGCGCGCTTGAAGGAATATCGGCTAC 270
726 rgThrThrArgThrSerGlyValAlaValIleGluGlyAsnAsnProSer 742
271 ATT...GTCCGCTTTCCGAT.....CA 290
743 LeuArgLeuLysAlaSerAspArgValValValAsnMetGlyAlaAlaHi 759
291 CGGCGCAAAATCCATCCGCTTC.....GACAAC.....C 322
759 sLysAsnGlnAlaTyrArgProLeuLeuLeuThrThrAspAsnGlyIleL 776
323 ATCCCTCACATTCGATTCGAGGACCGGCTAGTCCCGTTGACGGATTC 372
776 ysAlaTyrHisSerAspGlnGluAlaAlaGly.....786
373 AGCCTTTACCGCATCCATTTGGGACGATACGAA.....CACCATCCGCG 416
787 ...LeuValArgTyrThrAsnAspArgGlyGluLeuIlePheThrAlaAl 802
417 CGAC.....GGCTATGAGGGGCACAGCGCGGCTATCC.....453
802 aAspIleLysGlyTyrAlaAsnProGlnValSerGlyTyrLeuGlyValT 819
454 ..GCTCCCAAGCGCGGAGGATATATACAGCTACGCACATAAAGCGTT 501
819 rpValProValGlyAla.....AlaAla 826
502 GCCCAAAATATCCGCTCAACCTGACGACACCGCAGCAGCGGACACG 551
827 AspGlnAspValArgValAlaAlaSerThrAlaProSerThrAspGlyLy 843
552 GCTTCCGACCGTTTCCAAATGCCGGGCTATGCTGACGCAAGGAGTAG 601
843 s.....SerValHisGlnAsnAlaAlaLeuAspSerArgValMetP 857
602 GCGACGATTCACACCGCGCCATACGATACGCGCCGCTGACAGATCG 651
857 heGluGlyPheSerAsnPheGlnAlaPheAlaThrLysLysGluGluTyr 873
652 GCGAATGCGCGCGAGCGCTTCAAC.....GGCAC 680
874 ThrAsnValValIleAlaLysAsnValAspLysPheAlaGluTrpGlyVa 890
681 TGAGATATGTCAAAACATCATCGCGCGCGCAGAGAAATTTGCGCG 730
890 lThrAsp.....PheGluMetAlaProGlnTyrValSerS 902
731 CAGGCGATCCCGTG.....CAGGGTATAAGCAAGGCTCAAAATTT 771
902 erThrAspGlySerPheLeuAspSerValIleGlnAsnGlyTyrAlaPhe 918
772 GCTGTATGACGCGGCTTGGGCTGCTTCCACCGCAAAACAGATGGCGCG 821
919 ThrAspArgTyrAspLeuGly...IleSerLysProAsnLysTyrGlyTh 934
```

seq\_name: /cqn2\_6/ptodata/1/1aa/6A\_COMB.pep:US-08-783-774-2

seq\_documentation\_block:

; Sequence 2, Application US/08783774

; Patent No. 6054130

; GENERAL INFORMATION:

; APPLICANT: Spaete, Richard

; APPLICANT: Jackman, Winthrop

; TITLE OF INVENTION: NON-SPICING VARIANTS OF

; TITLE OF INVENTION: GP350/220

; NUMBER OF SEQUENCES: 19

; CORRESPONDENCE ADDRESS:

; ADDRESS: Pennie & Edmonds

; STREET: 1155 Avenue of the Americas

; CITY: New York

; STATE: NY

; COUNTRY: USA

; ZIP: 10036/2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

```

; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/783,774
; FILING DATE: 15-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Coruzzi, Laura A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7682-037
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-790-9090
; TELEFAX: 212-869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 907 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-783-774-2

```

```

alignment_scores:
  Quality: 122.50      Length: 412
  Ratio: 0.641        Gaps: 17
Percent Similarity: 46.359 Percent Identity: 22.330

alignment_block:
US-09-303-518D-463 x US-08-783-774-2 ..
Align seg 1/1 to: US-08-783-774-2 from: 1 to: 907

110 TGCACGCTCAGCATTTGCAACCCGACGAGGAATACCACTATTTCGGCAGC 159
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
469 SerThrAlaAspValThrSerProThrProAlaGlyThrThrSerGlyAl 485
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
160 AGGGGGAGCTGCCNAGCGCAACGCCCATATCGGATTCGGGAACATACA 209
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
485 aSerProValThrProSerPro.....SerProThrPasp..... 496
210 AAGCCATCAGTTGGGCCACCTGATTCACACAGCGCGCGTGTGAAGGAA 259
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
497 .....AsnGlyThrGluSerLysAla 503
260 ATATCGGCTACATGTCGCGCTTTTCGATCAGC.....GGCAC 297
|||:|||||:|||||:|||||:|||||:|||||:|||||:
504 ProAspMetThrSerSerThrSerProValThrThrProThrProAsnAl 520
298 AATTCCTATTCGCGCTTCGACACCATGCTCAGATTCGATTCGACGA 347
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
520 aThrSerProThrProAlaValThrThrPro...ThrProAsnAlaThrS 536
348 AGCGGCTAGTCGCGTTCGACGATTCAGCCTTTACCGCATCCATTGGGACG 397
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
536 erProThrProAlaValThrThrProThrProAsnAlaThrSerProThr 552
398 GATACGAACACATCCCGCGGAGGTATGACGGCCACAGGCGCGCGGC 447
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
553 LeuGlyLysThrSerProThrSerAlaValThr..... 563
448 TATCCCGCTCCCAAGGCG..... 466
564 .ThrProThrProAsnAlaThrSerProThrThrLeuGlyLysThrSerPro 580
467 ..CGAGGGATATATACAGTACG.....ACATAAAAGCGGTGCGCC 505
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
580 hrSerAlaValThrThrProThrProAsnAlaThrSerProThrLeuGly 596
506 AAATATCCGCTCAACCTACGACGACACCGACCGACCAACAGCGGCTT 555
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
597 LysThrSerProThrSerAlaValThrThrProThrProAsnAlaThrGl 613

```

```

556 GCGACGCTTTCCACAATGCCGGCGCTGCTGACGCAAGGAGTAGGCGA 605
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
613 yProThrValGly..... 617
606 CGGATTCAAACGGCCACCCGAGATACAGCCCGGAGCTGGACAGATCGGCA 655
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
618 .....GluThrSerProGlnAlaAsnAlaThrAsnHisThrLeuGlyGly 632
556 ATGCCGCGGAGAGCTTCACGGGCACTCCAGATATCGTCAAAACATCATC 705
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
633 ThrSerProThrProValValThrSerGlnProLysAsnAlaThrSerAl 649
706 GCGCGG.....CAGGAGAAATTG.....TCGGCGCAGCGCATGC 740
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
649 aValThrThrGlyGlnHisAsnIleThrSerSerSerThrSerSerMets 666
741 CGTCAGGGTATAAGCGAAGGCTCAACATTTGCTGTCATCAGCGGCTGG 790
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
666 erLeuArgProSerSerAsnProGluThrLeuSerProSerThrSer... 681
791 GTCGCTTTCCACCGAAACAAGATGGCGGCATCAACGATTTGGCAGAT 840
681 ..... 681
841 ATGGCGCACTCAAAAGACT...ATGCCGACGAGCCATCCGCGATT... 883
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
682 ....AspAsnSerThrSerHisMetProLeuLeuThrSerAlaHisProT 697
884 .....GGGCAGTCCAAACCCCAATGCCGACAGGCAATAGAACCGCTCA 928
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
697 hrGlyGlyGluAsnIleThrGlnValThrProAlaSerIleSerThrHis 713
929 GCATATCTTTATGGCAGCCATCCCATCAAGAGGATTTGGAGCTGTCGG 978
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
714 HisValSerThrSerSerProGluProArgProGly..... 725
979 GGAATATACGGCTTGGCGGCATCAGCGCACATCCTGTCAAGCGGTGCGCA 1028
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
726 ....ThrThrSerGlnAlaSerGlyProGlyAsnSerSerThrSerThrL 741
1029 GATGGCGCGATCCATTTGCCGAAAGGAAATCCGCGTCAGCGACAAATT 1078
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
741 ysProGlyGluValAsnValThrLysGlyThrProGlnAsnAlaThr 757
1079 TTGCGGATGGGCATACGCCAAATACCGCTCCCTTACCATTCCCGAAT 1128
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
758 SerPro.....GlnAlaProSerGlyGlnLysThrAlaValProThrVa 772
1129 ATCCGTTCAA.....ACTTGGACGAGCTTACGCAAGAAACATCAC 1172
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
772 lThrSerThrGlyGlyLysAlaAsnSerThrThrGlyGlyLysHisThrT 789
1173 CTCCTCAACCGTCCCGCGTCAACGCGCAAAATGTCAAACTGGCAGACC 1222
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
789 hrGlyHisGlyAlaArgThrSerThrGluProThrThrAspThrGlyGly 805
1223 AACGCCACCGCA...AGACAGCGCTACCGTTTGACG 1255
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
806 AspSerThrThrProArgProArgThrAsnAlaThr 817
seq_name: /cgn2_6/ptodata/1/iaa/pCTUS_COMB.pep:PCT-US95-04611A-19
seq_documentation_block:
; Sequence 19, Application PC/TUS9504611A
; GENERAL INFORMATION:
; APPLICANT: Spaete, Richard and Jackman, Winthrop, T.
; TITLE OF INVENTION: Non Splicing Variants of gp350/220
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooley Godward Castro Huddleson & Tatum
; STREET: 5 Palo Alto Square
; CITY: Palo Alto

```



```

; Sequence 6, Application PC/TUS9603916
; GENERAL INFORMATION:
; APPLICANT: Wild, Martha A.
; APPLICANT: Cochran, Mark D.
; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/03916
; FILING DATE: 23-MAR-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/126,597
; FILING DATE: 24-SEP-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39116-A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 985 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US96-03916-6

```

```

alignment_scores:
  Quality: 122.50      Length: 519
  Ratio: 0.554         Gaps: 26
  Percent Similarity: 42.582      Percent Identity: 22.351

```

```
alignment_block:

```

```
US-09-303-518D-463 x PCT-US96-03916-6 ..
```

```
Align seg 1/1 to: PCT-US96-03916-6 from: 1 to: 985
```

```

71 CCTCAGATTGGCAAGAGATCCCTTTATCC.....GGCAGGTT 108
   ||| ||| ||||| ||| ||| ||| ||| ||| ||| |||
366 ProAlaIleThrLeuIleProArgSerThrSerAspMetGlyPhep 382
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
109 CTGACCGTCAGATTTCGAACCCGCGGAAATACCACTATTCGGCG 158
   ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
382 eSerThr.....AlaArgAlaThrGlySerGluThrLeuSerVal. 395
159 CAGGGGGGAGCTTGCNAGCGCAACGCCCATATCGGATTGGGNAACATAC 208
395 .....
209 AAGCCATCAGTTGGGCCACCTGATGATTCACAGCGCGCGTTGAAGA 258
396 .....ProValGlnI 399
259 AATATCGGTACATTGTCGCTTTCCGATCAGC..... 292
399 uThrAspArgThrLeuSerThrThrProLeuThrProG 416
293 .....GGCACAAATTCATTGCGCCCTTCGACACCAACCATGCTCACATTCCG 337

```

```

416 lyGluSerGluAsnThrLeuPhePro...ThrThrAlaProGlyIleSer 431
   : : : : : : : : : : : : : : : : : : : : : : : : : :
338 ATTCTGACG.....AA 348
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
432 ThrGluThrProSerAlaAlaHisGluThrThrGlnThrGlnSerAlaGl 448
   : : : : : : : : : : : : : : : : : : : : : : : : : :
349 GCGGTAGTC.....CCGTGACGGATTCAGCC..... 376
   : : : : : : : : : : : : : : : : : : : : : : : : : :
448 uThrValValPheThrGlnSerProSerThrGluSerGluThrAlaArgS 465
   : : : : : : : : : : : : : : : : : : : : : : : : : :
377 .....TTTACCGCATCCATTGGGACGATACG 403
   : : : : : : : : : : : : : : : : : : : : : : : : : :
465 erGlnSerGlnGluProTyrPheThrGlnThrProSerThrGluGln 481
404 AACACCATCCCGCCGACGCTATGACGGCCACACGGCGCGCTATCCC 453
   : : : : : : : : : : : : : : : : : : : : : : : : : :
482 AlaAlaLeuThrGlnThrGlnIleAlaGluThrGluAlaLeuPheThrGl 498
   : : : : : : : : : : : : : : : : : : : : : : : : : :
454 GCTCCCAAGGCGGA.....GGGATATATA 479
   : : : : : : : : : : : : : : : : : : : : : : : : : :
498 nThrProSerAlaGluGlnMetThrPheThrGlnThrProGlyAlaGluT 515
480 CA.....GCTACGACATAAAGGCGTGCCTCAATAAATA..... 511
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
515 hrGluAlaProAlaGlnThrProSerThrIleProGluIlePheThrGln 531
512 TCCGCTCAACCTGACCGACACCGACGACCGGCTTCCGAC 561
   ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
532 SerArgSerThrProGluThrAlaArgAlaProSerAlaAlaProGl 548
562 CTTTTCCCAATGCGCGCGCTATGTGACGCAAGGAGTAGGCG...ACGG 608
   ||| ||||| : : : : : : : : : : : : : : : : : : : :
548 uValPheThrGlnSerSerSerThrValThrGluValPheThrGlnThrP 565
609 ATTCAACGCGCCACCGCATACAGCCGAGCTGGACAGATCGGCAATG 658
   ||| ||||| ||| ||||| ||| ||||| ||| ||||| ||| |||||
565 roSerThrValProLysThrThrLeuSerSerSerThrGluProAlaIle 581
659 CCGCGAAGCCTTCAACGGCA...CTGCAGATATCGTCAAAACATCATC 705
   : : : : : : : : : : : : : : : : : : : : : : : : : :
582 PheThrArgThrGlnSerAlaGlyThrGluAlaPheThrGlnThrSerSe 598
706 GCGCGCGCAGGAGAAATTGTCGCGCAGCGATGCCGTGACGGGTATAG 755
   ||||| : : : : : : : : : : : : : : : : : : : : : :
598 rAlaGluProAspThrMetArgThrGlnSerThrGluThrHisPhePhe 615
756 CGAAGCTCAAAACATTGTCTCATGCACGGCTTGGGTCTGCTTTCCACG 805
   : : : : : : : : : : : : : : : : : : : : : : : : : :
615 hrGlnAlaProSerThrVal.....Pro 622
806 AAAACAAGATGCGCGCATCAACGATTTGGCAGATATGCGCAACTCAA 855
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
623 LysAlaThrGlnThrProSerThrGluProGluValLeuThrGlnSerPr 639
856 GACT.....A 860
639 oSerThrGluProValProPheThrArgThrLeuGlyAlaGluProGluI 656
861 TGCCCGCAGCAGCCATCCCGGATTGGGCGAG.....TCCAAAACC 898
   : : : : : : : : : : : : : : : : : : : : : : : : : :
656 leThrGlnThrProSerAlaAlaProGluValTyrThrArgSerSerSer 672
899 CCAATGCGCGCACAGGCATAGACCGCTCAGCAATATCTTTATGGCACC 948
   ||||| : : : : : : : : : : : : : : : : : : : : : :
673 ThrMetProGluThrAla.....GlnSe 680
949 ATCCCATCAAAAGGATTGGAGCTGTCCGGGAAAATACGCGCTTGGCGG 998
   : : : : : : : : : : : : : : : : : : : : : : : : : :
680 rThrPro.....LeuA 684
999 CATCAGCGCACATCTGTCAAGCGGTCCAGATGGCGCGCATCGCATTCG 1048
   ||||| : : : : : : : : : : : : : : : : : : : : : :

```

```
684 laSerGluAsnProThrSerSerGlyThrGlyThrHisAsnThrGluPro 700
1049 CGAAGGGAATCCCGCGTACCGACAATTTTCGGCATCGCGCATACGCC 1098
701 Arg.....ThrTyrProValGlnThrThrPr 709
1099 AAATACC.....CGTCCCTTACCATTCGCCGAA 1127
709 ohisThrGlnLysLeuTyrThrGluAsnLysThrLeuSerPheProThrV 726
1128 TATCCGTTTCAAACTTGGACGCGTTACGGCAAGAAACATCACTCCT 1177
726 alValSerGluPheHisGluMetSerThrAlaGluSerGlnThrProLeu 742
1178 CAACCGTCCGCGGTCAACAGCGCAAAATGTCAAAC...TGGCAGACCAA 1224
743 LeuAspValLysIleValGluValLysPheSerAsnAspGlyGluValTh 759
1225 CGCACCCCGAAGACAGCGGTACCGTTTGACGGTAAAGGT...TTCCGAA 1271
759 rAlaThrCysValSer.....ThrValLysSerProTyrArgV 772
1272 TTTTGAGACGCGTGAATATGATACGAGCTCGATATCAAGAAATAT 1321
772 al..GluThrAsnTrpLysValAspLeu...ValAspValMetAspGluI 787
1322 CGGGGGCGGTATACCTAAGCTTAAGCTGTGTTTCATCGCAACCGAGA 1371
787 leSerGlyAsnSerProAlaGly.....ValPheAsnSerAsnGluLys 801
1372 TGGGAG 1377
802 TrpGln 803
```

seq\_name: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep: PCT-US96-03916-66

# seq\_documentation\_block:

; Sequence 66, Application PC/TUS9603916

; GENERAL INFORMATION:

; APPLICANT: Wild. Martha A.

; TITLE OF INVENTION: RECOMBINANT INFECTIOUS LARYNGOTRACHEITIS VIRUS

; TITLE OF INVENTION: AND USES THEREOF

; NUMBER OF SEQUENCES: 72

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham LLP

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10036

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: PCT/US96/03916

; FILING DATE: 23-MAR-1995

; CLASSIFICATION:

; PRIORITY APPLICATION DATA:

; APPLICATION NUMBER: US 08/126,597

; FILING DATE: 24-SEP-1993

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28 678

; REFERENCE/DOCKET NUMBER: 39116-A

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 66:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 985 amino acids

; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US96-03916-66

## alignment\_scores:

Quality: 122.50 Length: 519  
Ratio: 0.554 Gaps: 26  
Percent Similarity: 42.582 Percent Identity: 22.351

## alignment\_block:

US-09-303-518D-463 x PCT-US96-03916-66 ..

Align seg 1/1 to: PCT-US96-03916-66 from: 1 to: 985

```
71 CTCAGATTGGCAACGATCCCTTTATCC.....GGCAGGTT 108
||| ||| ||||| ||| ||||| ||| |||||
366 ProAlaIleThrThrLeuIleProArgSerThrSerAspMetGlyPhePh 382
109 CTCGACCGTCAGCATTTTCGAACCCGCGGAAATACCACTATTTCGGCAG 158
||||| ||||| ||||| ||||| ||||| |||||
382 eSerThr.....AlaArgAlaThrGlySerGluThrLeuSerVal. 395
159 CAGGGGGAGCTTCCCNAGCGCAACGCCCATATCGGATTGCGGAACATAC 208
395 ..... 395
209 AAGCCATCAGTTGGGCCACCTGATGATTCAACAGGGCGCGTTGAAGGA 258
||| ||||| ||||| ||||| ||||| |||||
396 .....ProValGlnG 399
259 AATATCGGTACATGTCGCGTTTCCTCCGATCACG..... 292
||| ||||| ||||| ||||| ||||| |||||
399 uThrAspArgThrLeuSerThrThrProLeuThrLeuProLeuThrProG 416
293 .....GGCACAAATTCATTCGCGCTTCGCAACCATGCGCTCACATCCG 337
||| ||||| ||||| ||||| ||||| |||||
416 lYgluSerGluAsnThrLeuPhePro...ThrThrAlaProGlyIleSer 431
338 ATCTGTGCG.....AA 348
||| ||||| ||||| ||||| ||||| |||||
432 ThrGluThrProSerAlaAlaHisGluThrThrGlnThrGlnSerAlaGl 448
349 CCGGTAGTC.....CCGTTGACGGATTTCAGCC..... 376
||| ||||| ||||| ||||| ||||| |||||
448 uThrValValPheThrGlnSerProSerThrGluSerGluThrAlaArgS 465
377 .....TTTACCGCATCCATTGGGACGGATACG 403
||| ||||| ||||| ||||| ||||| |||||
465 erGlnSerGlnGluProTrpTyrPheThrGlnThrProSerThrGluGln 481
404 AACACATCCCGCGCGGTATGACGGCCACAGGGCGCGGCTATCC 453
||| ||||| ||||| ||||| ||||| |||||
482 AlaAlaLeuThrGlnThrGlnIleAlaGluThrGluAlaLeuPheThrGl 498
454 GTCCTCCAAAGCGCGA.....GGGATATATA 479
||| ||||| ||||| ||||| ||||| |||||
498 nThrProSerAlaGluGlnMetThrPheThrGlnThrProGlyAlaGluT 515
480 CA.....GCTACGACATAAAAGCGTTGCCCAAAATA..... 511
||| ||||| ||||| ||||| ||||| |||||
515 hrGluAlaProAlaGlnThrProSerThrIleProGluIlePheThrGln 531
512 TCCGCTCAACGTGACCAACACCGCAGCAGCGGACACGCGCTTCCGCGAC 561
||| ||||| ||||| ||||| ||||| |||||
532 SerArgSerThrProProGluThrAlaArgAlaProSerAlaAlaProGl 548
562 CGTTTCCCAATGCGCGCGCTATGCTGACGCAAGGAGTAGCG...ACGG 608
||| ||||| ||||| ||||| ||||| |||||
548 uValPheThrGlnSerSerThrValThrGluValPheThrGlnThrP 565
609 ATTCAAGCGCGCCACCCGATACAGCCCGAGCTGGACAGATCGGGCAATG 658
```







```
374 GCCTTACCGCATCATTTGGGAGGATACGACACCATCCCG..... 415
      ::::::::::: ||| ::::::::::: |||
202 ProMetThrThrLeuProThr...ThrArgThrLeuProThrThrTh 217
      ::::::::::: ||| ::::::::::: |||
416 .....CCGACGGCTATGACGGGCCACAGGCGCGCTATCCGCTCC 458
      ::::::::::: ||| ::::::::::: |||
217 rThrLeuProThrThrMetThr.....LeuP 226
      ::::::::::: ||| ::::::::::: |||
459 CAAAGCGCGGAGGATATACAGCTACGACATAAAAGCGCTGGCCAAA 508
      ::::::::::: ||| ::::::::::: |||
226 roMetThrThrThrLeuProThrThrThr.....LeuPro.... 238
      ::::::::::: ||| ::::::::::: |||
509 ATATCCGCTCAACCTGACGACACCGACGACGCGGACGCTGCC 558
      ::::::::::: ||| ::::::::::: |||
239 ...ThrThrThrLeuProThrThrThrLeuProThrMetThrLeuP 254
      ::::::::::: ||| ::::::::::: |||
559 GACCGTTCCCAATGCGCGGCTATGCTGACGCAAGGAGTAGGCGAGG 608
      ::::::::::: ||| ::::::::::: |||
254 oThrThrThrThrLeuProThrMet.....MetThrLeuProThr 268
      ::::::::::: ||| ::::::::::: |||
609 ATTCAACGCGCCCGCATACAGCC...CCGAGCTGGACA.....GAT 649
      ::::::::::: ||| ::::::::::: |||
268 hrThrThrLeuProThrThrThrLeuProThrThrMetValSer 284
      ::::::::::: ||| ::::::::::: |||
650 CGGGCAATGCCCGGAGCCTTCAACGCGCTGACATATCGTCAAAAC 699
      ::::::::::: ||| ::::::::::: |||
285 ThrPheValProProThrPro...LeuProMetGlnAsnHisGluProVa 300
      ::::::::::: ||| ::::::::::: |||
700 ATCATCGCGCGCGGAGGAGAAATGTCGGCGGACGCGGATCGGTCAGGG 749
      ::::::::::: ||| ::::::::::: |||
300 lalaThrSerProSerSerProGlnProAlaGluThrHisProVal.... 315
      ::::::::::: ||| ::::::::::: |||
750 TATAAGCGAAGCTCAACATGCTGTCATGCGCGCTGGTCTGCTTT 799
      ::::::::::: ||| ::::::::::: |||
316 .....ThrLeuLeuGlyAlaThr..... 321
      ::::::::::: ||| ::::::::::: |||
800 CCACCGAAACAAAGATGGCGCGCATCAACGATTTGGCAG..... 838
      ::::::::::: ||| ::::::::::: |||
322 .....ArgThrGlnProThrSerSerProLeuTyrSerTyrThrThrAs 336
      ::::::::::: ||| ::::::::::: |||
839 .....ATATGGCGCACTCAA 854
      ::::::::::: ||| ::::::::::: |||
336 pGlySerAspThrValThrGluSerSerAspGlyLeuTrpAsnAsnAng 353
      ::::::::::: ||| ::::::::::: |||
855 AGACT 859
      ::::::::::: ||| ::::::::::: |||
353 lnThr 354
      ::::::::::: ||| ::::::::::: |||

seq_name: /cgn2_6/ptodata/1/1aa/5A_comb.pep:US-08-314-309A-6

seq_documentation_block:
; Sequence 6, Application US/08314309A
; Patent No. 5677141
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: FUKAGAWA, MASAO
; APPLICANT: IWAMI, MORITA
; APPLICANT: ARAMORI, ICHIRO
; APPLICANT: KOJO, HITOSHI
; TITLE OF INVENTION: PROCESS FOR PRODUCING 7-AMINOCEPHEM
; TITLE OF INVENTION: COMPOUND OR SALTS THEREOF
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/314,309A
; FILING DATE: 30-SEP-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/631,906
; FILING DATE: 21-DEC-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5677141man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 18-863-0 CONT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-314-309A-6

alignment_scores:
      Quality: 120.00      Length: 324
      Ratio: 0.745      Gaps: 15
      Percent Similarity: 49.691      Percent Identity: 22.222

alignment_block:
US-09-303-518D-463 x US-08-314-309A-6 ..
Align seg 1/1 to: US-08-314-309A-6 from: 1 to: 402

76 GATTGTGCAACAGATCCCTTTATCCGCGCAGGTCTTCGACCGTCAGCATTT 125
      ::::::::::: ||| ::::::::::: |||
90 GluIleLysAsnAsnProAspValLeuGluVal..... 100
      ::::::::::: ||| ::::::::::: |||
126 CGAACCCGCGGGAAATACCACTATTTCGCGCAGCAGGGGGAG..... 168
      ::::::::::: ||| ::::::::::: |||
101 .GluGluAspGlnLeuTrpHisLeuPheAspGluGlnAspGluGlyGluP 117
      ::::::::::: ||| ::::::::::: |||
169 .....CTTGCNAGCGCAACGCCATATCGCGCAGCAGGGGGAG..... 168
      ::::::::::: ||| ::::::::::: |||
117 heSerThrAlaAlaLeuValThrGlnAsnGlyAlaTrpGlyLeuGlyThr 133
      ::::::::::: ||| ::::::::::: |||
205 ATACAAAGCCATCAGTTGGGC.....CACTGTGATGATTCACAGCGGC 248
      ::::::::::: ||| ::::::::::: |||
134 IleSerHisArgGlnProGlySerThrSerTyrIleTyrAspSerAl 150
      ::::::::::: ||| ::::::::::: |||
249 CGTTGAAGGAATATCGCTACATTTCCGCTTTTCGATCACGGG...C 295
      ::::::::::: ||| ::::::::::: |||
150 aGlySerGlyThrTyrAlaTyrValVal.....AspThrGlyIleL 164
      ::::::::::: ||| ::::::::::: |||
296 ACAAAATCCATTCCCTTCGACAAACCATCCCATTCGATTCGTGAC 345
      ::::::::::: ||| ::::::::::: |||
164 euGluSerHisAsnGluPheSerGlyArgAlaIleThrGlyTyrAsnAla 180
      ::::::::::: ||| ::::::::::: |||
346 GAACCGGTAGTCCGTTGACGGATTTCAGCCTTTACCGCATCCATTGGGA 395
      ::::::::::: ||| ::::::::::: |||
181 ValGlyGlySerAsnAlaAsp.....ThrAs 189
      ::::::::::: ||| ::::::::::: |||
396 CGGATACGACACCATCCCGCGCAGCGCTATGACGGCGCACAGCGCGC. 444
      ::::::::::: ||| ::::::::::: |||
189 nGlyHisGlyThrHisValAla.....GlyThrIleGlyAla 202
      ::::::::::: ||| ::::::::::: |||
445 ..GGCTATCCCGCTCCCAAGAGCGCGAGG..... 471
      ::::::::::: ||| ::::::::::: |||
202 rgThrTyrGlyValAlaLysAsnThrAsnLeuIleAlaValLysValPhe 218
      ::::::::::: ||| ::::::::::: |||
472 .....GATATATACACTACGA 488
      ::::::::::: ||| ::::::::::: |||
```



```
274 aValAsnAspIleIleAsnArgGlyArgGlnAsnLysAlaAlaIleSerM 291
536 GCAGCAGCGGACAAAGCGTTGCCGACCGTTCACCAATGCC..... 576
111 111 111 111 111 111 111 111 111 111 111 111
291 etSerLeuGlyGlyGlyTyrSerAlaPheAsnAsnAlaValAsnThr 307
577 .....GGCGCTATGCTGACGACGAGTAGGCGGAGTTCACAA 614
111 111 111 111 111 111 111 111 111 111 111 111
308 AlaTyrSerArgGlyValLeuSerValValAlaAlaGlyAsnAspAsnGl 324
615 ACGCGCCACCGCATACAGCCCGGAGCTGCGAGATCGGGCAATCGCCGCG 664
111 111 111 111 111 111 111 111 111 111 111 111
324 nAsnAlaAlaAsnTyrSerPro.....AlaSerAlaAlaA 336
665 AAGCC..... 669
111 111 111 111 111 111 111 111 111 111 111 111
336 snAlaIleThrValGlySerIleAlaSerAsnTrpAlaArgSerSerPhe 352
670 .....TTCAACGGCACTGCAGATATCGTCAAAACATCATCGCGCGG 713
111 111 111 111 111 111 111 111 111 111 111 111
353 SerAsnTyrGlySerValLeuAspIlePheAlaProGlyThrSerIleLe 369
714 AGGAGAAATGTCGGCGAGCGAGTCCGTCGAGCGGTATAAGCGAGGCT 763
111 111 111 111 111 111 111 111 111 111 111 111
369 uSerAlaTrpIleGlyGlyAsnSerAlaThrAsnThrIleSer...Gly 385
764 CAACATTGCTCTCATGCGCGCTGGGCTG..... 795
111 111 111 111 111 111 111 111 111 111 111 111
385 hrSerMetAlaThrProHisValThrGlyValValLeuTyrLeuGlnAla 401
796 .....CTTCCACCGCAAAACAGATGGCGGCATCAACGATTGGC 836
111 111 111 111 111 111 111 111 111 111 111 111
402 LeuGluGlyLeuThrThrSerGlyAlaAlaAlaArgLeuAsnAlaLeuAl 418
837 AGATATGGCGCAACTCAAGAC 858
111 111 111 111 111 111 111 111 111 111 111 111
418 aThrThrGlyArgValSerAsn 425
seq_name: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.US-08-728-323A-2
seq_documentation_block:
; Sequence 2, Application US/08728323A
; Patent No. 5948676
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A.
; APPLICANT: Russo, James J.
; APPLICANT: Edelman, Isidore S.
; APPLICANT: Moore, Patrick S.
; TITLE OF INVENTION: Immediate Early Protein From Kaposi's
; TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA
; TITLE OF INVENTION: Encoding Same And Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/728,323A
; FILING DATE:
; APPLICATION NUMBER:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/52268/JPW/MSC/SKS
```

```
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-278-0400
; TELEFAX: 212-391-0525
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1162 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-728-323A-2

alignment_scores:
Quality: 120.00 Length: 336
Ratio: 0.876 Gaps: 16
Percent Similarity: 40.774 Percent Identity: 23.810

alignment_block:
US-09-303-518D-463 x US-08-728-323A-2 ..
Align seg 1/1 to: US-08-728-323A-2 from: 1 to: 1162
69 CGCCTCAGATTGGCAACGATCCCTTTATCCGGCAGGTTCTCGACCGTC 118
111 111 111 111 111 111 111 111 111 111 111 111
7 ArgLeuArgSerGlyArgSerThrGlyAlaProLeu...ThrArgGly 22
119 AGCATTTTCGACCGCGAGGAAATACCACTATTTCGGCAGCAGGGGAG 168
111 111 111 111 111 111 111 111 111 111 111 111
22 rCysArgLysArgAsnArgSer.....ProGluArgCysAspLeuGlyA 37
169 CTTGCCNAGCGCAACGCGCCATATCGG..... 194
111 111 111 111 111 111 111 111 111 111 111 111
37 spAspLeuHisLeuGlnProArgArgLysHisValAlaAspSerIleAsp 53
195 .....ATTGGAAACATACAAACCATCAGTTGGGGCCACCTGATGA 235
111 111 111 111 111 111 111 111 111 111 111 111
54 GlyArgGluCysGlyProHisThrLeuProIleProGlySerProThrVa 70
236 TTCACAGCGCGCGTGTGAAGGAAATATCGCTACATTGTCGCTTTTC 285
111 111 111 111 111 111 111 111 111 111 111 111
70 lPheThrSerGly.....LeuProAlaPheV 79
286 GATCAGCGGCACAAATTCATTTCGCCCTTCGACAAACCATCGCTCAT 335
111 111 111 111 111 111 111 111 111 111 111 111
79 alSerSerProThrLeuProValAla..... 87
336 CGATTCTGACGACCGCGTAGTCCCGTTGACGGATTGACGCTTTTACG 385
111 111 111 111 111 111 111 111 111 111 111 111
88 .....ProfileProSe 91
386 TCCATTGGGACGGATACGACACCATCCCGCGACGGCTATGACGGCCA 435
111 111 111 111 111 111 111 111 111 111 111 111
91 rProAlaProAlaThrProLeuPro...ProAlaLeuLeuProProVa 107
436 CAGGCGCGCGCTATCCCGCTCCCAAAGCGCGAGGATATATACAGCTA 485
111 111 111 111 111 111 111 111 111 111 111 111
107 lThrThrSerSerSerProIleProProSerHisProValSerProGly 124
486 CGACATAAAGCGGTGCCCCAAAATATCCCGCTCAACCTGACCCACAAC 535
111 111 111 111 111 111 111 111 111 111 111 111
124 hrThr.....AspThrHisSerProSerPro 132
536 GCAGCAGCGGACAAAGCGTTGCCGACCGTTCACCAATGCGCGGCTA.. 583
111 111 111 111 111 111 111 111 111 111 111 111
133 AlaLeuProProThrGlnSerProGlnSerSerGlnArgProProLeuSe 149
584 .....TGCTGACGCAAGGAGTAGGCGACGAGATTCAACGCGCCACC 625
111 111 111 111 111 111 111 111 111 111 111 111
149 rSerProThrGlyArgProAspSerSerThrProMetArgProProPro 166
626 .....GATACAGCCCGCGAGCTGGACAGATCGGGCAATCGCCCGCA 670
111 111 111 111 111 111 111 111 111 111 111 111
166 erGlnGlnThrThrProHisSerProThrThrProProProGluPro 182
```



[illegible]



```

;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: US/08/697,221
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/004,802
; FILING DATE: 22-SEP-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Kodroff, Cathy A.
; REGISTRATION NUMBER: 33,980
; REFERENCE/DOCKET NUMBER: WST64AUSA
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-540-9206
; TELEFAX: 215-540-5818
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 393 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-697-221-16

```

```

alignment_scores:
  Quality: 117.50      Length: 451
  Ratio: 0.721        Gaps: 17
  Percent Similarity: 36.142  Percent Identity: 19.290

```

```

alignment_block:
US-09-303-518D-463 x US-08-697-221-16

```

```

Align seg 1/1 to: US-08-697-221-16 from: 1 to: 393
516 CCTCAACCTGACCGACACCGCAGCAGCGGACCGGTCGCGGACGGTT 565
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
58 ProGlyProaspGluAlaProArgMetProGluAlaAla..ProProVal 73
566 TCCACATGCGCGCGCTATGCTGACGCAAGAGTAGCGCGAGATTCAAA 615
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
74 AlaProAlaPro.....AlaAlaProThrPr 82
616 CGCGCCACCGATACACGCCGAGCTGGACAGATCGGCGCAATCGCGCGA 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
82 oAlaAlaProAlaProAlaProSerTrpProLeuSerSerValProS 99
666 AG..... 667
99 erGlnLysThrTyrGlnGlySerTyrGlyPheArgLeuGlyPheLeuHis 115
668 CCTTCACGGCAGTGCAGATATCGTCAAAACATCATCGCGCGGAGGA 717
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
116 SerGlyThrAlaLysSerValThrCysThrTyrSerProAlaLeuAsnly 132
718 GAAATTTCGCGCGCAGCGGATCGCGTCAGCGGATTAAGCGAAGCTCAA 767
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
132 sMetPheCysGlnLeuAlaLysThrCysProVal..... 143
768 CATTTGTCATGACCGGCTTGGTTC.....TGCTTTCCACGAAACA 811
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
144 .....GlnLeuTrpValAspSerThrProProGlyThr 155
812 AGATGGCGCGCATCAACGATTTGGCAG..... 838
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
156 ArgValArgAlaMetAlaIleTyrLysGlnSerGlnHisMetThrGluVa 172
838 ..... 838

```

```

172 lValArgArgCysProHisHisGluArgCysSerAspSerAspGlyLeuA 189
838 ..... 838
189 laProProGlnHisLeuIleArgValGluGlyAsnLeuArgValGluTyr 205
839 .....ATATGGCGCAACTCAAGACTATGCCGCGCAGCAGCC 873
206 LeuAspArgArgAsnThrPheArgHisSerValValProTyrGluPr 222
874 ATCCGGGATTGGCAGTCCCAAAACCCCAATGCGCGCACAAAGGCATAGAAGC 923
222 o.ProGluValGlySer.....AspCysThrThrIleHis..... 233
924 CGTCAGCAATATCTTTATGGCAGCCATCCCATCAAAAGGATTTGGAGCTG 973
234 ...TyrAsnTyrMetCysAsnSer.....Sercy 242
974 TCCGGGGGAAATACGG ..... 989
242 sMetGlyGlyMetAsnArgArgProIleLeuThrLeuGluA 259
990 .....CTTGGCGGCATCACGGCACATCTCTCAAGCG 1022
259 spSerSerGlyAsnLeuLeuGlyArgAsnSerPheGluValCysValCys 275
1023 GTCCGAGATGGCGCGATCGCATTCGCGAAAGGAAATCCGCCGTCAGCG 1072
276 AlaCysProGlyArgAspArgArgGluGluGluAsnLeuArgLys... 291
1073 ACAATTTTCCGCGATCGGCATACGCCAATACCCGTCCTCCCTTACCATTCC 1122
291 ..... 291
1123 CGAAATATCCGTTCAAACTTGGAGCAGCGTTACGGCAAGAAACATCAC 1172
292 .....LysGlyGluProHisHis 297
1173 CTCCTCAACCGTCCCGCGTCAACGGCAAAATGTCAAACCTGGCAGACC 1222
298 GluLeuProProGlySerThrLysArgAlaLeuProAsnAsnThrSerSe 314
1223 AACGCCACCCGAGAGCAGCGGTACCGTTTACGCGTAAAGGGTTTCCGAAT 1272
314 rSerProGlnProLysLysLys..ProLeuAspGlyGluTyrPhe..... 328
1273 TTTGAGAACACGTGAAATATGATACGAGCTCGATATTCAAGAATATATC 1322
329 .....ThrLeuGlnIleArg..... 333
1323 GGGGGCGGTATACCTAAGGCTAAGCCTGTGTTGATCGGAACCGAGAT 1372
334 .....GlyArgGluArgp 338
1373 GGGAGGTTGATAGGAAGCTTAATAAATTGACAACCTCGTGAGCAGGTGGAG 1422
338 heGluMetPheArgGluLeuAsnGluAlaLeuGluLeuLysAspAlaGln 354
1423 AAAAATGTTACGAAACGAGAGAGAGTAGTACAGTAGTCAAGTTTAAAGC 1472
355 AlaGlyLysGluProGlyGlySerArgAlaHisSerSerHisLeuLys... 370
1473 CCATCCGCAACGAGAGTGGGAAATAAACAGGGTTAGATTTTATCAT 1522
370 ..... 370
1523 TTATAGTGGTGATATCAATAAGAAAGGCACAGTAACAGAGGCGCATAGT 1572
371 .....SerLysLysGlyGlnSerThrSerArgHisLys 381
1573 CTAACCCCGTGTGATGTACGGGTGATACAAACAAACCTCGGCACCTGTAT 1620
382 .....LysLeuMetPheLysThrGluGlyProAsp 391

```

```

alignment_scores:      Quality: 117.00      Length: 684
                       Ratio:  0.369      Gaps:    34
Percent Similarity: 46.345  Percent Identity: 18.713

alignment_block:
US-09-303-518D-463 x US-08-296-791-5  ..

Align seg 1/1 to: US-08-296-791-5 from: 1 to: 1702

      88 GATCCCTTTATCCGCCAGTTCTC..... 111
      ||| |||::: ::::|||
171 AspLysPheValThrGluValAlaProIleGluAlaSerThrAlaSer 187
      ||| :||| ::::: ||| |||
112 .GACCGTCAGCATTTCCGAACCCGACGGGAATACCCATTATC..... 153
      ||| :||| ::::: ||| |||
187 rAspAlaGlyThrTyraSnAspGlnAsnLysTyProAlaPheValArgL 204
154 .....GGCAGCAGGGGGGAGCTT 171
204 euGlySerGlyThrGlnPheIleTyrLysGlyAspAsnTySerLeu 220
172 GCNAGCGCACGGCCATATCCGGATTCGGAAACATACAAAGCCATCAGTT 221
      ::::: ||| ::::: ||| ::|||:::
221 IleLeuAsnAsnHisGluValAlcIcLYaSnAsnLeuLys..... 233
222 GGGCCACCTGATGATTCAACAGCGCGCGGTGAAGGAAATATCGGC.... 267
234 .....LeuValGlyAspAlaTyThrTyrGlyLeuAlaGlyThrP 274

```



```
1070 GGCACAAATTTCCGATCGCGCATACGCCAAATACCCCTCCCTTAC... 1116
      |||
507 alLeuAsn.....AspAspLysGlnValAspProAsnSerIleTyrPhe 521
      |||
1116 ..... 1116

522 GlyPheArgGlyGlyArgLeuAspAlaAsnGlyAsnAsnLeuThrPheG1 538
1117 .CATTCGCCGAATATC..... 1131
      |||
538 uHisIleArgAsnIleAspAspGlyAlaArgLeuValAsnHisAsnThrS' 555
1132 ..CGTTCAAACTGGAGCAGCGTTACGGCAAGAAACATCACCTCC... 1176
      |||
555 erLysThrSerThrValThrIleThrGlyGluSerLeuIleThrAspPro 571
1177 TCACCGGTCGCCCGCTCAACCGGCAAAATGTCAAACCTGGCAGACCAACG 1226
      |||
572 AsnThrIleThrPro.....TyrAsnIleAspAlaProAspGluAs 585
1227 CCACCGC.....AAGACAGCGCTACCGTTTGACGCGTA 1258
      |||
585 pAsnProTyrAlaPheArgArgIleLysAspGly.....GlyG 598
1259 AAGGTTTCCGAATTTTGAGAGCACGTCGAATAT..... 1293
      |||
598 InLeuTyrLeuAsnLeuGluAsnTyrThrTyrTyrAlaLeuArgLysGly 614
1294 ...GATACGAAGCTCGATATTCAGAATATTCGGGGCGGTATACCTAA 1340
      |||
615 AlaSerThrArgSerGluLeuProLysAsnSerGly..... 626
1341 GCCTAAGCCTGTGTTGATGCCAACCGAGTGGAGGTTGATAGGAAC 1390
      |||
627 .....GluSerAsnGluAsnTrp.....L 633
1391 TTAATAAATTCACAACTCGTGAGCAGGTGGAGAAAATGTTCAGGAAACG 1440
      |||
633 euTyrMetGlyLysThrSerAspAlaAlaLysArgAsnValMetAsnHis 649
1441 AGAAGAAGAGTCAGAGTAGTCAGTTTAAAGCCCAT...GGCAACAGAGA 1487
      |||
650 IleAsnAsnGluArgMetAsnGlyPheAsnGlyTyrPheGlyGluGluG1 666
1488 ATGGGAAATAAACAAGCGGTAGATTTTAAATCATTTTATAGTGGTGATA 1537
      |||
666 uGlyLysAsnAsnGlyAsnLeuAsnValThr...PheLysGlyLysSerG 682
1538 TCAATAAGAAAGGCACAGTACAGGAGGCGCATAGTCTAACCCGCTGGTAT 1587
      |||
682 LuGlnAsnArgPheLeuLeuThrGlyGlyThrAsnLeu...AsnGlyAsp 697
1588 GTACGGGTGATACACAACAC.....TCGGCACCTGATAAACA 1625
      |||
698 LeuLysValGluLysGlyThrLeuPheLeuSerGlyArgProThrProHi 714
1626 TGGGGTTTATCAAGCCACAGTGGAAATTAAGCCTGAT..... 1665
      |||
714 sAlaArgAspIleAlaGlyIleSerSerThrLysLysAspGlnHisPheA 731
1666 .....GGAAAGTTGGGAGGTGAAACAGAAA 1689
      |||
731 laGluAsnAsnGluValValGluAspAspTrpIleAsnArgAsnPhe 747
1690 AAGGTGGGAAAGTG...ATGACCAAGCACACCATGTTCCCAAAAGATTG 1736
      |||
748 LysAlaThrAsnIleAsnValThrAsnAsnAlaThrLeuTyrSerGlyAr 764
1737 GGATAGGCTAGAAATTAGGCTTCAAGTACTTCGGCTTGGGAAAGTAGAA 1786
      |||
764 gAsnValAlaAsnIleThrSerAsnIleThrAlaSerAspAsnAlaLysV 781
1787 TA 1788
```

```
::
781 al 781
```

```
seq_name: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:PCT-US95-10661A-5
```

```
seq_documentation_block:
; Sequence 5, Application PC/TUS9510661A
; GENERAL INFORMATION:
; APPLICANT: Washington University, et al.
; TITLE OF INVENTION: Haemophilus Adherence and Penetration Protein
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Flehr, Hohnbach, Test, Albritton & Herbert
; STREET: 4 Embarcadero Center, Suite 3400
; CITY: San Francisco
; STATE: California
; COUNTRY: United States
; ZIP: 94111-4187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/10661A
; FILING DATE: 16-AUG-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/296,791
; FILING DATE: 25-AUG-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Trecartin, Richard F.
; REGISTRATION NUMBER: 31,801
; REFERENCE/DOCKET NUMBER: FP-59941/RFT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 781-1989
; TELEFAX: (415) 398-3249
; TELEX: 910 277299
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1702 amino acids
; TYPE: amino acid
; TOPOLOGY: unknown
PCT-US95-10661A-5
```

```
alignment_scores:
Quality: 117.00 Length: 684
Ratio: 0.369 Gaps: 34
Percent Similarity: 46.345 Percent Identity: 18.713
```

```
alignment_block:
```

```
US-09-303-518D-463 x PCT-US95-10661A-5 ..
Align seg 1/1 to: PCT-US95-10661A-5 from: 1 to: 1702

88 GATCCCTTTATCCGCGAGGTTCTC..... 111
||| |||: |||
171 AspLysPheValThrGluValAlaProIleGluAlaSerThrAlaSerSe 187
112 .GACCGTCAGCATTTCCGAACCGCAGCGGAAATACCACCTATTC..... 153
||| |||: |||
187 rAspAlaGlyThrTyrAsnAspGlnAsnLysTyrProAlaPheValArgL 204
154 .....GCACGAGGGGGAGCTT 171
|||: |||
204 euGlySerGlyThrGlnPheIleTyrLysLysGlyAspAsnTyrSerLeu 220
172 GCCNAGCGCAACGGCCCATATCGGATTTCGGAAACATACAAACCATCAGTT 221
|||: |||
221 IleLeuAsnAsnHisGluValGlyGlyAsnAsnLeuLys..... 233
```







```
164 GGGAGCTTCCNAGCGCAACGCCATATCGGATGGGAAACATACAAAGC 213
    :::::|||||:::
490 ThrAlaAlaProThrAlaAla
214 CATCAGTTGGGCCCTGATGATTCACAGCGCGCGTTGAAGGAAATAT 263
    |||
497 .....ProThrAlaAlaProT 502
264 CGGTACATGTCGCTTTTCGATCAGCGGCACAAATTCATTCGCGCT 313
    :::::|||||:::
502 ThrAlaAlaProThrAlaAlaProSerThrValAlaProProAlaThrPro:518
    :::::|||||:::
314 TCGACAACCATCCCTCACCATCCGATTCGACGAAGCGGTAGTCCCGTT 363
    :::::|||||:::
519 ProAlaThrAlaAlaProValProProThrThrAlaAlaProProAl 535
    :::::|||||:::
364 GACGATTCAGCTTACCGATCCGATTCGACGAGCAGC 397
    :::::|||||:::
535 aProThrAlaAlaProThrAlaAlaProThrThrAlaAlaProGluSerP 552
    :::::|||||:::
398 ..GATACGAACACCATCCGCGCGAGCGGTATG...ACGGGCCACAGCGC 442
    ||| |||:::|||||:::
552 roThrThrValThrValProProThrThrAlaAlaProThrThrAlaAlaProThr 568
    ||| |||:::|||||:::
443 CGCGTATCCCG...CTCCCAAGCGCGAGGATATATACAGCTACGAC 489
    :::::|||||:::
569 ThrAlaValProGluilePro.....IleThrValThrSe 580
    :::::|||||:::
490 ATAAAGCGGTGCGCAAAATATCCGCTCAACCTGACCGGACACCGCG 539
    :::::|||||:::
580 rAlaProThrAlaAlaProThrThrAlaAlaProThrThrAlaAlaProThrAlaA 597
    ||||| :::::|||||:::
540 CACCGGACAAAGCGTTCGCGACCGTTTCCACAATGCCGCGCTATGCTGA 589
    ||||| :::::|||||:::
597 laProThrThrAlaValProGluile..... 605
    :::::|||||:::
590 CCAAGGAGTAGGCGACGGATTCAAACCGCCGACCCGATACAGCCCGGAG 639
    ||| :::::|||||:::
606 .....ProThrThrValThrSerProProThrAlaAlaProTh 618
    :::::|||||:::
640 CTGGACAGATCGGCAATGCCCGCGAGCGCTCAACGCGCAGCTGCAGATA, 688
    : :::::|||||:::
618 r.....ThrAlaAlaProAlaProAlaProAlaProAlaProThValThrValP 631
    :::::|||||:::
689 .....TCGTCAAAACATCATCGCGCGCGAGGAGAAATTCGCGCG 730
    :::::|||||:::
631 roProThrAlaAlaProThrThrAlaAlaProAlaProAlaProAlaProAla 647
    :::::|||||:::
731 CAGCGGATCGCGTGCAGGGTATAGCGAAGGCTCAACATTCGCTGTCATG 780
    ||| ||| :::::|||||:::
648 ThrAlaProProThrAlaAlaProThrThrAlaAlaProAlaProAlaProAla 664
    :::::|||||:::
781 CACGCTTGGTCTCTTCCACCAACAAACAGATGGCGCGCATCAACGA 830
    ||| ||| :::::|||||:::
664 rThr.....ValThrValProProThrAlaAlaProThrAlaAla.... 677
    :::::|||||:::
831 TTGGCAGATATGGCGCACTCAAGACTATGCGCGCAGCAGCCATCGCG 880
    ||||| :::::|||||:::
678 .....ProProThrValAlaAlaHisAlaProAlaProAla 687
    :::::|||||:::
881 ATTGGGAGTCCAAACCCCAATGCCGCAAGGATAGAGCGGTGAGC 930
    :::::|||||:::
688 ThrAlaAlaProValThrThrThrSerAlaProAlaThrThrProGluAs 704
    :::::|||||:::
931 ATATCTTTATGGCAGCATCCCATCAAGGATTCGAGTGTCCGGGG 980
    :::::|||||:::
704 pAspAspileAspProProLeuProLeuProAlaHisAlaAspCysAspLysTrp 721
    :::::|||||:::
981 AAAAT.....ACGCTTGGGG.....GCATCAGCGCACATC 1012
    ||:::|||||
721 luGluCysAsnValLeuProThrAlaHisAlaAspCysAspLysTrp 737
    ||:::
1013 CTGTCAAGCGGTGCGAGATGGCGCGATCGCATTCGCCGAAAGGA..... 1057
```

```

    :::::|||||:::
738 ValCysAspGlyAsnAsnGlnValLeuValValCysSerGluGlyLeuG1 754
    ||||| :::::|||||
1058 .....AATCCCGCTCAGCGACA 1075
    ||||| :::::|||||
754 nPheAsnProThrThrLysThr 761
```

seq\_name: /cgn2\_6/ptdata/1/laa/6A\_COMB.pep:US-09-017-706-9

seq\_documentation\_block:

; Sequence 9, Application US/09017706A

; Patent No. 6087147

; GENERAL INFORMATION:

; APPLICANT: ITO, YOSHIMU

; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING

; TITLE OF INVENTION: MALTOPENTAOSE, VECTOR CONTAINING SAID GENE AND

; FILE REFERENCE: 8361-0003-0

; CURRENT APPLICATION NUMBER: US/09/017,706A

; CURRENT FILING DATE: 1998-02-05

; EARLIER APPLICATION NUMBER: JP 305071/1997

; EARLIER FILING DATE: 1997-10-21

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9

; LENGTH: 614

; TYPE: PRT

; ORGANISM: Pseudomonas sp., Strain KO-8940

; US-09-017-706-9

alignment\_scores:

Quality: 115.50 Length: 291

Ratio: 0.868 Gaps: 16

Percent Similarity: 45.704 Percent Identity: 27.491

alignment\_block:

US-09-303-518D-463 x US-09-017-706-9 ..

Align seg 1/1 to: US-09-017-706-9 from: 1 to: 614

287 ATCAAGGGACAAATTCATTCGCGCTTCGACAAACCATGCTCCATCC 336

349 MetSerAlaThrAlaSerThrArgAlaAlaAlaThrProAlaThrAl 365

337 GATTCGTACCAAGCGGTAGTCCCGTTGACGGATTGAGCTTTACCGCAT 386

365 aArgThr.....ThrAlaThrAlaArgProGlyG 376

387 CCATTGGGACGATACGACACCATCCGCGCGACGCTATG..... 427

376 lArgGlyThrGlyThr.....ProAlaArgValLeuGlnPro 389

428 .....ACGGGCCACAGGCGCGGTATCCCGCTCCCAAGGC 465

390 AspArgGlyTrpLeuGlyLeuArgAlaProLeuAlaArg...HisArg 405

466 GCAGGGGATATATACAGTACGACAT.....AAAGGCGT 500

405 InHisGlyGlyLeuProGlnArgHisArgGlyGlnLeuValArgGlnArg 421

501 TGCCCCAAATATCCGCTCAACCTGACCGCAACACCGCAGCGGACAC 550

422 LeuValGluGlnArgGlnGlnProAspArgLeu...ArgProArgG1 437

551 GGCTTGGCGCGCTTTCCCAATTCGCGCGCTATGCTGACGCAAGAGTA 600

437 nGlyLeuArgArgHisGlnGlnGlyArgTyrGlyAlaAspAlaSerA 454

601 GCGCAGGATTCAAACGCGCGCCCGCATACAGCCCGAGCTGGACAGATC 650

454 rg.....ProAlaCysArgPro 459

651 GGCAATGCGCGGAGAGCCTTCAAGCGCACTGCAGATATCGTCAAAAACA 700  
||| |||||:||||:||||| ||| ||||| ||  
460 GlyValLeuArgArgAspLeuArgArg...LeuArgGlnArgGlnLeuHi 475  
701 TCATCGCGCGGAGAGAAATGTCGGCGC..... 731  
||||:||||| ||| |||||:||||| ||| ||||| ||  
475 sArgHisGlyGlyAspArgGlyGlyArgProArgAspAlaValArgT 492  
732 ..AGGGGATCGCGTGCAGGGTATAAGCGAGGCTCAAACTGCTGTCAT 779  
||| |||||:||||| ||||| ||| ||||| ||  
492 hrArgLeuTrpArgGlyGlyAspProArgArgCysAlaHisArgHis 508  
780 GCACGG.....CTTGGGTGCTTTCCACCGGAAACAAAGATGG 817  
||| |||||:||||:||||:||||:||||| ||| ||||| ||  
509 AlaAlaGlyAlaValLeuSerLeuThrPheAsnGluThrAlaAspTh 525  
818 CGCGCATCAACGA.....TTTGGCAGATATGCGCGCACTCAA 855  
||||:||||| ||| |||||:||||| ||| ||||| ||  
525 rValTrpGlyGlnAsnLeuPheValValGlyAsnValGlyAla.LeuGly 541  
856 GACTAT.....GCCGACGAGCCTCCCGGATGGCGAGTCCAAAACC 899  
||||:||||:||||| ||| ||||| ||| ||||| ||  
542 AsnTrpAlaProAlaAlaGlyAlaAlaMetThrTrpIleSerGlySerG 558  
900 CAATCGCGCACAGGCATAGAACCGCTCAGCAATATCTTTATGGCAGCA 949  
||||:||||| ||| |||||:||||| ||| ||||| ||  
558 ySerThrGlyGln..... 567  
950 TCCCCATCAAGGGATGGAGCTGTCGCGGGGAAATACGGCTTGGCGGC 999  
||| |||||:||||| ||||| ||| ||||| ||  
567 alGlnLeuProAlaAspThrProValGlnTrpLysTyr..... 579  
1000 ATCAGCGCACATCCTGCTCAAGCGGTGCGAGATGGCGCGCATGCGC 1049  
||||:||||| ||| |||||:||||| ||| ||||| ||  
580 .....ValLysLysAspGlyAlaGlyAsnValValTrpG 591  
1050 GAAAGGAAATCGCGGTC 1068  
||||:||||| ||| ||||| ||  
591 uSerGlyGlyAsnArgVal 597

seq\_name: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep.us-09-017-706-10

seq\_documentation\_block:  
; Sequence 10, Application us/09017706A  
; Patent No. 6087147  
; GENERAL INFORMATION:  
; APPLICANT: ITO, YOSHIFUMI  
; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING  
; TITLE OF INVENTION: MALTOPEPTASE, VECTOR CONTAINING SAID GENE AND  
; TITLE OF INVENTION: TRANSFORMANT  
; FILE REFERENCE: 8361-0003-0  
; CURRENT APPLICATION NUMBER: US/09/017.706A  
; EARLIER FILING DATE: 1998-02-05  
; EARLIER APPLICATION NUMBER: JP 305071/1997  
; EARLIER FILING DATE: 1997-10-21  
; NUMBER OF SEQ ID NOS: 14  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 10  
; LENGTH: 614  
; TYPE: PRT  
; ORGANISM: Pseudomonas sp., Strain KO-8940  
US-09-017-706-10

alignment\_scores:  
Quality: 115.50 Length: 291  
Ratio: 0.868 Gaps: 16  
Percent Similarity: 45.704 Percent Identity: 27.491

alignment\_block:  
US-09-303-518D-463 x US-09-017-706-10 ..  
Align seg 1/1 to: US-09-017-706-10 from: 1 to: 614

287 ATCAGGGGCACAAATTCATTGCGCTTCGACAAACCATGCTCCATTC 336  
||||:||||| ||| ||| ||||:||||| ||| ||||| ||  
349 MetSerAlaThrAlaSerThrArgAlaAlaAlaThrThrProAlaThrAl 365  
337 GATTCTCAGCAAGCGGTAGTCCCGTTGACGGATTACGCTTTACCGCAT 386  
||| ||||:||||| ||| ||||| ||| ||||| ||  
365 aArgArgThr.....ThrAlaThrAlaArgProGlyG 376  
387 CCATTGGGAGGATACGAACACCATCCGCGGAGGCTATG..... 427  
||| ||||| ||| |||||:||||| ||| ||||| ||  
376 lYArgGlyThrGlyThr.....ProAlaArgValLeuGlnPro 389  
428 .....ACGGCCACAGGGCGCGGCTATCCCGTCCCAAGGC 465  
||| ||||| ||| |||||:||||| ||| ||||| ||  
390 AspArgGlyTrpLeuGlyLeuArgAlaPro.LeuAlaArg...HisArg 405  
466 GCGAGGATATATACAGCTAGCAT.....AAAAGCGT 500  
||||:||||| ||| |||||:||||| ||| ||||| ||  
405 lnhHisGlyGlyLeuProGlnArgHisArgGlyGlnLeuValArgGlnArg 421  
501 TGCCCAAAATATCGGCTCAACCTGACCGACAAACCGCAGCACCGGACAAC 550  
||| ||||| ||| |||||:||||| ||| ||||| ||  
422 LeuValGlnGlnArgGlnGlnProAspArgLeu...ArgProArgArgG 437  
551 GCCTTGGCGACCGTTTCCACAATCCCGCGCTATGTCACGCAAGGAGTA 600  
||| ||||| ||| |||||:||||| ||| ||||| ||  
437 nGlyLeuArgArgHisGlnGlnGlyArgTyrglyAlaAspAlaAlaSerA 454  
601 GCGCAGGATTCAAACCGCGCCACCGGATACAGCCCGAGCTGGACAGATC 650  
|| ||||| ||| |||||:||||| ||| ||||| ||  
454 rg.....ProAlaCysArgPro 459  
651 GGGCAATCGCGCGGAGCGCTTCAAGCGCACTGCAGATATCGTCAAAAACA 700  
||| |||||:||||| ||||| ||| ||||| ||  
460 GlyValLeuArgArgAspLeuArgArg...LeuArgGlnArgGlnLeuHi 475  
701 TCATCGCGCGGAGGAGAAATGTCGGCGC..... 731  
||||:||||| ||| |||||:||||| ||| ||||| ||  
475 sArgHisGlyGlyAspArgGlyGlyCysArgArgProArgAspAlaValArgT 492  
732 ..AGCGATGCGGTGCGAGGTATAGGGAAGGCTCAAACTGCTGTCAT 779  
||| |||||:||||| ||||| ||| ||||| ||  
492 hrArgLeuTrpArgGlyGlyAspProArgArgCysAlaHisArgArgHis 508  
780 GCACGG.....CTTGGGTGCTTTCCACCGGAAACAAAGATGG 817  
||| ||||| ||| |||||:||||| ||| ||||| ||  
509 AlaAlaGlyAlaAlaValLeuSerLeuThrPheAsnGluThrAlaAspTh 525  
818 CGCGCATCAACGA.....TTTGGCAGATATGCGCGCACTCAA 855  
||| ||||| ||| |||||:||||| ||| ||||| ||  
525 rValTrpGlyGlnAsnLeuPheValValGlyAsnValGlyAla.LeuGly 541  
856 GACTAT.....GCCGACGAGCCTCCCGGATGGCGAGTCCAAAACC 899  
||||:||||| ||| |||||:||||| ||| ||||| ||  
542 AsnTrpAlaProAlaAlaGlyAlaAlaMetThrTrpIleSerGlySerG 558  
900 CAATCGCGCACAGGCATAGAACCGCTCAGCAATATCTTTATGGCAGCA 949  
||||:||||| ||| |||||:||||| ||| ||||| ||  
558 ySerThrGlyGln.....TipArgAlaThrV 567  
950 TCCCCATCAAGGGATGGAGCTGTCGCGGGGAAATACGGCTTGGCGGC 999  
||| |||||:||||| ||||| ||| ||||| ||  
567 alGlnLeuProAlaAspThrProValGlnTrpLysTyr..... 579  
1000 ATCAGCGCACATCCTGCTCAAGCGGTGCGAGATGGCGCGCATGCGC 1049  
||||:||||| ||| |||||:||||| ||| ||||| ||  
580 .....ValLysLysAspGlyAlaGlyAsnValValTrpG 591  
1050 GAAAGGAAATCGCGGTC 1068  
||||:||||| ||| ||||| ||  
591 uSerGlyGlyAsnArgVal 597

seq\_name: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap:US-09-017-706-11

seq\_documentation\_block:  
; Sequence 11, Application US/09017706A  
; Patent No. 6087147

GENERAL INFORMATION:  
; APPLICANT: ITO, YOSHIFUMI

; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING  
; TITLE OF INVENTION: MALTOSE, VECTOR CONTAINING SAID GENE AND

; TITLE OF INVENTION: TRANSFORMANT

; FILE REFERENCE: 8361-0003-0

; CURRENT APPLICATION NUMBER: US/09/017,706A

; EARLIER FILING DATE: 1998-02-05

; EARLIER FILING DATE: 1997-10-21

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 11

; LENGTH: 614

; TYPE: PRT

; ORGANISM: Pseudomonas sp., Strain KO-8940

US-09-017-706-11

alignment\_scores:

Quality: 115.50 Length: 291

Ratio: 0.868 Gaps: 16

Percent Similarity: 45.704 Percent Identity: 27.491

alignment\_block:

US-09-303-518d-463 x US-09-017-706-11

Align seg 1/1 to: US-09-017-706-11 from: 1 to: 614

287 ATACGGGCGACAAATCCATTCGCGCTCGACACCATCGCTCACATTCC 336

349 MetSerAlaThrAlaSerThrArgAlaAlaThrProAlaThrAl 365

337 GATTCTGACGAAGCGGTAGTCCCTGACGGATTACGCTTTACCGCAT 386

365 argArgThr.....ThrAlaThrAlaArgProGlyG 376

387 CCATTGGGACGGATACGACACCATCCGCGCGCTATG..... 427

376 lyArgGlyThrGlyThr.....ProProAlaArgValLeuGlnPro 389

428 .....ACGGGCCACAGGGCGGCTATCCGCTCCCAAGGC 465

390 AspArgGlyTrpLeuGlyLeuAlaPro..LeuAlaArg...HisArg 405

466 GCGAGGATATACAGCTACGACAT.....AAAAGCGT 500

405 INHISGlyLeuProGlnArgHisArgGlyGlnLeuValArgGlnArg 421

501 TGCCCAAAATATCCGCTCAACCTGACGCGCAACCGCGACCGACAAC 550

422 LeuValGlnGlnArgGlnProAspArgLeu...ArgProArgGln 437

551 GCGTTCGCGACCGTTCCAAATGCGCGCTATGCTGACGCAAGGAGTA 600

437 nGlyLeuArgArgHisGlnGlnArgTrpGlyAlaAspAlaAlaSerA 454

601 GCGAGGATTCAACCGCGCCCGATACACCGCCGAGCTGGACAGATC 650

454 rg.....ProAlaCysArgPro 459

651 GGGCAATGCCGCGAAGCCTTCAAGCGCACTGCAGATATCGTCAAAACA 700

460 GlyValLeuArgArgAspLeuArg...LeuArgGlnArgGlnLeuHi 475

701 TCATCGCGCGCGAGAGAAATTTGCGGCGC..... 731

475 sArgHisGlyGlyAspArgGlyCysArgArgProArgAspAlaValArg 492

732 ..AGGCGATCGGTGACGGGTATAGCGAAGGCTCAAAACATTGCTGTCT 779  
492 hrArgLeuTrpArgGlyGlyAspArgArgCysAlaHisArgArgHis 508  
780 GCACGG.....CTTGGGTCTGCTTCCACGCAAAACAAGATGG 817  
509 AlaAlaGlyAlaAlaValLeuSerLeuThrPheAsnGlnThrAlaAsp 525  
818 CGCGCATCAACGA.....TTTGCAGATATGGCGCAACTCAAA 855  
525 rValTrpGlyGlnAsnLeuPheValValGlyAsnValGlyAla..Leu 541  
856 GACTAT.....GCGCAGCAGCATCCGCGATTGGCGCAGTCCAAAAC 899  
542 AsnTrpAlaProAlaAlaGlyAlaAlaMetThrTrpIleSerGlySer 558  
900 CAATGCGCGCAAGGCGATAGACCGCTCAGCAATATCTTTATGGCAG 949  
558 ySerThrGlyGln.....TrpArgAlaThrV 567  
950 TCCCATCAAGGATTGGAGCTGTCGGGGGAAATACGGCTTGGGCGGC 999  
567 alGlnLeuProAlaAspThrProValGlnTrpLysTyr..... 579  
1000 ATCAGCGCACATCTGTCAAGCGGTGCGAGTGGCGCGCATCGCATTC 1049  
580 .....ValLysLysAspGlyAlaGlyAsnValValTrpG 591  
1050 GAAAGGAAATCCGCGCTC 1068  
591 uSerGlyGlyAsnArgVal 597

seq\_name: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pap:US-09-017-706-12

seq\_documentation\_block:

; Sequence 12, Application US/09017706A

; Patent No. 6087147

GENERAL INFORMATION:

; APPLICANT: ITO, YOSHIFUMI

; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING

; TITLE OF INVENTION: MALTOSE, VECTOR CONTAINING SAID GENE AND

; TITLE OF INVENTION: TRANSFORMANT

; FILE REFERENCE: 8361-0003-0

; CURRENT APPLICATION NUMBER: US/09/017,706A

; EARLIER FILING DATE: 1998-02-05

; EARLIER FILING DATE: 1997-10-21

; NUMBER OF SEQ ID NOS: 14

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 12

; LENGTH: 614

; TYPE: PRT

; ORGANISM: Pseudomonas sp., Strain KO-8940

US-09-017-706-12

alignment\_scores:

Quality: 115.50 Length: 291

Ratio: 0.868 Gaps: 16

Percent Similarity: 45.704 Percent Identity: 27.491

alignment\_block:

US-09-303-518d-463 x US-09-017-706-12

Align seg 1/1 to: US-09-017-706-12 from: 1 to: 614

287 ATACGGGCGACAAATCCATTCGCGCTCGACACCATCGCTCACATTCC 336

349 MetSerAlaThrAlaSerThrArgAlaAlaThrProAlaThrAl 365

337 GATTCTGACGAAGCGGTAGTCCCGTTCAGCGATTACGCTTTACCGCAT 386

```

365 aArgArgThr.....ThraThraAlaArgProGlyG 376
387 CCATTGGGAGGATACGAACACCATCCGCGCAGCGGTATG..... 427
   :: ||||| ||| |||||::: ::
376 lyArgGlyThrGlyThr.....ProProAlaArgValLeuGlnPro 389
428 .....ACGGGCCACAGGCGCGGCTATCCGCTCCCAAGGC 465
   ||| ||||| |||||::: |||
390 AspArgGlyTrpLeuGlyLeuArgAlaPro.LeuAlaArg...HisArg 405
466 GCGAGGGATATACAGCTACGACAT.....AAAGCGGT 500
   ::|||::: |||||::: |||||
405 InHisGlyLeuProGlnArgHisArgGlyGlnLeuValArgGlnArg 421
501 TGCCCCAAATATCCGCTCAACCTGACCGCAACCGCAGCACCGGACAAC 550
   ::: ||||| |||||::: |||
422 LeuValGlnArgGlnGlnProAspArgLeu...ArgProArgArgG 437
551 GCGTTCGCGAGGTTTCCAAATGCGCGGCTGTGCTGCAAGAGGATTA 600
   ::: ||| ||| ||||| |||||
437 nGlyLeuArgArgHisGlnGlnGlyArgTyrGlyAlaAspAlaAlaSerA 454
601 GCGCAGCGATTCAACGCGCCACCGATACAGCCCGGAGCTGGACATC 650
454 rg.....ProAlaCysArgPro 459
651 GGGCAATGCGCGGAAGCCTTCAACGCGCAGCTGACGATATCGTCAAAACA 700
   ||| |||||::: |||||::: |||
460 GlyValLeuArgAspLeuArgarg...LeuArgGlnArgGlnLeuHis 475
701 TCATCGCGCGGCGAGAAATTTGCGCGC..... 731
475 sArgHisGlyGlyAspArgGlyCysArgArgProArgAspAlaValArg 492
732 ..AGCGGATCGCGTCGACGGTATAGCCAGGCTCAACATTCGTGTCAT 779
   ||| ||||| ||||| |||||
492 hrArgLeuTrpArgGlyGlyAspProArgArgCysAlaHisArgArgHis 508
780 GCACGG.....CTTGGGTCTGCTTCCACCGCAAAACAAGATGG 817
   ||| |||||::: |||||::: |||
509 AlaAlaGlyAlaAlaValLeuSerLeuThrPheAsnGluThrAlaAspTh 525
818 CGGCGATCAACGA.....TTTGGCAGATATGGCGCAACTCAA 855
   ::: :: |||||::: |||||
525 rValTrpGlyGlnAsnLeuPheValValGlyAsnValGlyAla.LeuGly 541
856 GACTAT.....GCCGACGACCCATCCCGATTTGGCGAGTCCAAACCC 899
   ::|||::: ||||| |||||
542 AsnTrpAlaProAlaAlaGlyAlaAlaMetThrTrpIleSerGlySerG 558
900 CAATGCGGCACAAAGCATAGAAGCGCTCAGCAATATCTTTATGCGACCA 949
   ::|||::: |||||::: |||
558 ySerThrGlyGln.....TrpArgAlaThrV 567
950 TCCCCATCAAGGGATTTGAGCTGTCCGGGAAATACGGCTTGGCGCGC 999
   ::: :: |||||::: |||||
567 alGlnLeuProAlaAspThrProValGlnTyrLysTyr..... 579
1000 ATCAGCGCACATCTGTCTAAGCGGTGCGAGATGGCGGATCGCATGGC 1049
   |||||::: |||||::: |||
580 .....ValLysLysAspGlyAlaGlyAsnValValTrpG 591
1050 GAAAGGGAATCCGCGTC 1068
   ::|||::: |||
591 userGlyGlyAsnArgVal 597

seq_name: /cgn2_6/ptodata/1/1aa/6A_COMB.pep:US-09-017-706-13
seq_documentation_block:
; Sequence 13, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI

```

```

; TITLE OF INVENTION: A-AMYLASE GENE HAVING ABILITY FOR HIGHLY PRODUCING
; TITLE OF INVENTION: MALTOPEPTIDASE, VECTOR CONTAINING SAID GENE AND
; TITLE OF INVENTION: TRANSFORMANT
; FILE REFERENCE: 8361-0003-0
; CURRENT APPLICATION NUMBER: US/09/017,706A
; CURRENT FILING DATE: 1998-02-05
; EARLIER APPLICATION NUMBER: JP 305071/1997
; EARLIER FILING DATE: 1997-10-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 13
; LENGTH: 614
; TYPE: PRT
; ORGANISM: Pseudomonas sp., Strain KO-8940
US-09-017-706-13

alignment_scores:
    Quality: 115.50      Length: 291
    Ratio: 0.868        Gaps: 16
    Percent Similarity: 45.704    Percent Identity: 27.491

```

```
seq_name: /cgn2.6/ptodata/1/iaa/6A_COMB.pep:US-09-017-706-13
seq_documentation_block:
; Sequence 13, Application US/09017706A
; Patent No. 6087147
; GENERAL INFORMATION:
; APPLICANT: ITO, YOSHIFUMI
```



287 ATCAGGCGCAAAATTCATTGCCCTTCGCACCAACCATTGCCTCACATTCC 336  
::: ::::: ||| ||| :||| :||| |||  
349 MetSerAlaThrAlaSerThrArgAlaAlaAlaThrProAlaThrAl 365  
::: ::::: ||| ||| :||| :||| |||  
337 GATTCTGAGAACGGGTAGTCCCGTTGACGGATTACGCTTTACCGCAT 386  
||| ||| :||| :||| |||  
365 aArgArgThr..... ThrAlaThrAlaArgProGlyG 376  
:  
387 CCATTGGAGCGGATACGAACACCATCCCGCGGCGGTATG..... 427  
::: ||||| ||| ||||| :||| :|||  
376 IyArgGlyThrGlythr..... ProProAlaIargValLeuGlnPro 389  
:  
428 ..... ACGGGCCACAGGGCGGGGGTATCCGCTCCCAAGGC 465

```

seq_name: /cgn2_6/ptodata/1/1aa/6B_COMB pep: US-09-377-155-33

seq_documentation_block:
; Sequence 33, Application US/09377155
; Patent No. 6197312
; GENERAL INFORMATION:
; APPLICANT: PEAK, Ian Richard Anselm
; APPLICANT: JENNINGS, Michael Paul
; APPLICANT: MOXON, E. Richard
; TITLE OF INVENTION: NOVEL SURFACE ANTIGEN
; FILE REFERENCE: 065064/0128
; CURRENT APPLICATION NUMBER: US/09/377,155
; CURRENT FILING DATE: 1999-08-19
; PRIOR APPLICATION NUMBER: PC7/AU98/01031

```











```
1288 AAATATGATACGAGCTCGATATTCAGAAATATCG..... 1323
    : : : : : ||| : : : : |||
533 SerIleGlnAsnAspTyrAspGluAspTyrSerTyrTyrArgPr 549
1324 .GGGGCGGTATACCTAGCTGTAAGCTGCTGTTGATCGCAACGAGAT 1372
    ||| : : : : : ||| : : : : |||
549 oArgArgProIleProGlnGlyAspLeuTyrTyrLysAsnTyrArgT 566
1373 GGGAGGTTGATAGGAGCTTAATAATTGACA..... 1404
    : : : : : ||| : : : : |||
566 yTyrAlaLeuLysSerGlyArgLeuAsnAlaProMetProGluAsn 582
1405 ..... 1415
583 GlyValAlaGluAsnAsnAspTrpIlePheMetGlyTyrThrGlnGluG1 599
1416 GTGGAGAAAAATGTT.....CAGGAACGAGAGAGGAGTACAGA 1456
    : : : : : ||| : : : : |||
599 uAlaArgLysAsnAlaMetAsnHisLysAsnAsnArgIleGlyAspG 616
1457 GTAGTCAGTTT.....AAAGCCATCGCCCAACGAGAA 1488
    : : : : : ||| : : : : |||
616 luGlyGlyPhePheAspGluGluAsnGlyLysGlyHis..... 628,
1489 TGGGAAATATAACAGGTTAGATTTT.....AATCA 1520
    : : : : : ||| : : : : |||
629 ...AsnGlyAlaLeuAsnLeuAsnAsnGlyLysSerAlaGlnAsnAr 644,
1521 TTTTATA.....GGTGGTATATCAATAAGAAAGGCACAGTACAG 1561
    : : : : : ||| : : : : |||
644 gPheLeuLeuThrGlyAlaAsnLeuAsnGlyLysIleSerValThrG 661
1562 GAGGCATAGTCTACCCGTGT.....GATGACGGGTGATA 1599
    ||| : : : : |||
661 luGlyAsnValLeuLeuSerGlyArgProThrProHisAlaArgAspPhe 677
1600 CAACAAACCTCGCACCTGATAAATGGGTTTATCAAGCGACAGTGGGA 1649
    : : : : : ||| : : : : |||
678 ValAsnLysSerSerAlaArgLysAspAlaHisPheSerLysAsnG1 694
1650 AATTAAAGCCTGATGAAGTGGAGGTGAAACGAAAGAGTGGGA 1699
    : : : : : ||| : : : : |||
694 uValValPheGluAsp...AspTrpIleAsnArgThrPheLysAlaAlaG 710
1700 AAGTG...ATGACCAAGCACACCATGTCCCAAAAGATTGGATCAGGCT 1746
    : : : : : ||| : : : : |||
710 luIleAlaValAsnGlnSerAlaSerPheSerSerGlyArgAsnValAsp 726
1747 AGAATTAGGCTGAAGTTACTTCGGCTTGGGAAAGTAGAATAATGCTT 1794
    ||| : : : : : ||| : : : : |||
727 AspIleThrAlaAsnIleThrAlaThrAspAsnAlaLysValAsnLeu 742

seq_name: /cgn2_6/ptodata/1/iaa/5B_COMB.pep.us-08-224-482-4

seq_documentation_block:
; Sequence 4, Application US/08224482
; Patent No. 5837692
; GENERAL INFORMATION:
; APPLICANT: Mercola, Dan
; TITLE OF INVENTION: Inhibition of the Mitogenic Activity of
; TITLE OF INVENTION: PDGF by Mammalian EGR
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell and Flores
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92122
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
```

```
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/224,482
FILING DATE: 07-APR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-ME 9913
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 543 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-224-482-4

alignment_scores:
Quality: 113.50 Length: 445
Ratio: 0.631 Gaps: 23
Percent Similarity: 40.449 Percent Identity: 22.247

alignment_block:
US-09-303-518D-463 x US-08-224-482-4 ..
Align seg 1/1 to: US-08-224-482-4 from: 1 to: 543
89 ATCCCTTTATCCGCGAGTTCTCGACCGTCAGCATTCGAACCCGACGGG 138
||| : : : : ||| : : : : |||
131 IleThrTyrThrGlyArgPheSerLeuGluProAlaProAsn...SerG1 146
139 AAATACCACCTAT..... 151
||| : : : : |||
146 yAsnThrLeuTrpProGluProLeuPheSerLeuValSerGlyLeuValS 163
152 .....TCGGCAGCAGGGGGAGCTTGCNACGCGCAACGGCC 187
||| : : : : ||| : : : : |||
163 erMetThrAsnProAlaSerSerSerAlaProSerProAlaAla 179
188 ATATCGATTGGGNAACATACAA...GCCATCAGTTGGCCACCTCATG 234
||| : : : : ||| : : : : |||
180 SerSerAlaSerAlaSerGlnSerProLeuSerCysAlaVal..... 194
235 ATTCACAGCGCGCGTTGAAGGAAATATCGGCTACATTTCGCGCTTTTC 284
||| : : : : ||| : : : : |||
195 .....ProSerAsnAspSerSerProIleTyrSerAlaAlaP 207
285 CGATCAGCGGCACAAATTCATTTCGCGCTTCGACAAACCATGCTCACAT 334
||| : : : : ||| : : : : |||
207 roThrPheProThr.....ProAsnThr 214
335 CCGATTCTGACGAAGCGGTAGTCCCGTTACGCGATTACGCTTTACCGC 384
||| : : : : ||| : : : : |||
215 AspIlePheProGluPro.....GlnSerGlnAlaPheProG1 227
385 ATCCATTGGGACGATACGAACACCATCCCG.....CCGACGCGTA 425
||| : : : : ||| : : : : |||
227 ySerAlaGlyThrAlaLeuGlnTyrProProAlaTyrProAlaAlaL 244
426 TGACGGGCGCACAGGGCGGGCTATCCCG.....CTCCCAAAGGCG 466
||| : : : : ||| : : : : |||
244 ysGlyGlyPheGlnValProMetIleProAspTyrLeuPhePro.GlnG1 260
467 CGAGGGATATATA..... 479
||| : : : : |||
260 nGlnGlyAspLeuGlyLeuGlyThrProAspGlnLysProPheGlnGlyL 277
479 ..... 479
```

```
277 euGlusSerArgThrGlnGlnProSerLeuThrProLeuSerThrIleLys 293
479 .....
294 AlaPheAlaThrGlnSerGlySerGlnAspLeuLysAlaLeuAsnThrSe 310
480 .....CAGTACGACATAAAGGGTGGCCCAAAATATCGGCCTC 519
310 rTyrGlnSerGlnLeuLysProSerArgMetArgLysTyrProAsnA 327
520 AACTGACCGACACACCGCAGCAC...CGACAAACGGCTTGGCAGCGTTT 566
327 rgProSerLysThrProHisGlnArgProTyrAlaCys...ProVal 342
567 CCACAATGC...CGGGCTATGCTGACGCAAGGAGTAGCGCAGGATTCA 613
343 GluSerCysAspArgarghe.....SerArgSerAspGluLe 355
614 AACGCCACCCCGATACAGCCCGAGCTGGACAGATCGGGCAATCGCGC 663
355 uThrArgHis...IleArgIleHisThrGlyGln.....L 366
664 GAAGCTTCACCGGCACTGACATATCGTCAAAACATCATCGGCGCGC 713
366 ysProPheGln.....CysArgIleCysMetArgAsnPheSerArgSer 380
714 AGCAGAAATTGTGGCGCGAGGCGTCCGCTGAGGATATACGGAAGGCT 763
381 AspHisLeuThrThrHisIleArgThrHisThrGlyGluLysProPheAl 397
764 CAAACATTGCTGTATGACGCGCTGGGTCTGCTTTCCACCGAAACAAAG 813
397 acysAspIleCysGlyArgLysPheAlaArgSerAspGluArgLys.Arg 413
814 ATGGCGGATCAACCAATTGGCAGATATGGCGCAACTCAAGACTATGC 863
414 HisThrLysIleHis.....LeuArgGlnLysAspLysLys 425
864 CGCAGCAGCCATCCGGATGGGCAATCCAAACCCCAATCGCGCACAAAG 913
425 sAla.....AspLysSerValValAlaLaserSerAlaThrSerS 438
914 GCATAGAAGCGGTACACATATCTTTATGGCAGCCATCCCATCAAGGG 963
438 erLeuSerSerTyrProSerProValAlaThrSerTyrPro..... 451
964 ATTGGAGCTGTCGGGAAATACGGCTTGGGCGCATCACGCGCATGCC 1013
452 ...SerProValThrThrSerTyrProSerProAlaThrThrSerTyrPr 467
1014 TGTCAAGCGGTGCGAGATGGGCGGATCGCATTCGCGAAAGGAAATCCG 1063
467 o.....SerP 469
1064 CGGTACGCGACAAATTTGGCGATGGCGATACGCCAAATACCGTCCCT 1113
469 roValProThrSerPheSerProGlySerSerThrTyrProSerPro 485
1114 TACCATTCGGAAATATCGCTTCAACTTGGAGCAGCGTTCAGGCAAGA 1163
486 ValHisSerGlyPheProSerProSerValAlaThrThrTyr..... 499
1164 AACATCACCTCTCAACCGTTCGCGCGTCA 1194
500 .....SerSerValProProAla 505
```

seq\_name: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:us-09-165-239A-4

seq\_documentation\_block:

; Sequence 4, Application US/09165239A

; Patent No. 6344554

; GENERAL INFORMATION:

```
; APPLICANT: JOHNSON, ALEXANDER
; APPLICANT: BRAUN, BURKHARD R
; TITLE OF INVENTION: POLYNUCLEOTIDE SEQUENCES FROM CANDIDA
; TITLE OF INVENTION: ALBICANS ENCODING POLYPEPTIDES ASSOCIATED WITH FILAMENTOUS
; TITLE OF INVENTION: GROWTH
; FILE REFERENCE: 220022000700
; CURRENT APPLICATION NUMBER: US/09/165,239A
; CURRENT FILING DATE: 1998-10-01
; PRIOR APPLICATION NUMBER: 60/068,065
; PRIOR FILING DATE: 1997-12-18
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FASTSEQ for Windows version 3.0
; SEQ ID NO 4
; LENGTH: 750
; TYPE: PRT
; ORGANISM: Candida albicans
; US-09-165-239A-4
```

#### alignment\_scores:

Quality: 113.00 Length: 427  
Ratio: 0.653 Gaps: 18  
Percent Similarity: 40.515 Percent Identity: 20.141

#### alignment\_block:

US-09-303-518D-463 x US-09-165-239A-4 ..

Align seg 1/1 to: US-09-165-239A-4 from: 1 to: 750

```
29 TTCGTCCACTACTGGCAGTGTGCTGCGGATGATGCACACGCTCAGAT 78
||||| ||||| |||
266 PheCys.....TrpAspCysGlnCys.....AspThrPro..... 275
79 TTGCAAAACGATCCCTTTATCCGCGAGGTTCGACCGTCAGCATTTTCA 128
||| ||| ||||| |||
276 .....SerProSerThrThrThrValSerThrSerS 286
129 ACCGACGCGGAATACCACTATTTCGCGACGAGGGGAGCTTGCCNAGC 178
: : ||||| ||||| |||||
286 erAlaProSerThrSerProGluSerSerAla.....ProSer 298
179 GCAACGCCCATATCGGATGGGAACATACAAAGCCATCAGTTGGGCCAC 228
: : ||||| |||||
299 ThrThrThrVal..... 302
229 CTGATGATTCACACGCGCGGTGAAGAAATATCGGCTACATGTCCG 278
||| |||
303 .....ThrThrSerS 306
279 CTTTTCGATCAGGGGACAAATTCATTCGCCCTTCGACAAACCATGCCT 328
: : ||||| ||||| |||||
306 erSerProValThrSerProGluSerSerValProGluThrThrThr... 321
329 CACATTCGGATTCTCAGAAAGCCGCTAGTCCGTTGACGGATTACGCTT 378
321 ..... 321
379 TACCGCATCCATTGGGACGGATACGAACACATCCCGCGACGGCTATGA 428
||| ||| ||||| |||||
322 .....ValThrThrSerSerValProGluThr.....T 331
429 CGGCCACAGGGCGGCTATCCCGTCCCAAGCGCGAGGATATAT 478
|| : : ||||| ||||| |||||
331 hrProGluSerSerAla.....ProGluThrThrThrValThr 343
479 ACAGCTAGACATAAAGGGCTTGCCTCCAAATATCCGCC...TCAACGTG 525
|| : : ||||| ||||| |||||
344 ThrSerSerValProSerThrThrProGluSerSerAlaProGluThrTh 360
526 ACCGACAAACCGCAGCGGACAAACGCTTGCAGCGCTTCCACAATGC 575
||| : : ||||| ||||| |||||
360 rProGluSerSerAlaProGluSerSerValPro..... 371
```



```
576 CGGCGCTATCTGACGCAAGAGTAGGCGCAGGATTCAAACGCGCCACC 625
372 .....GlusSerAlaProGlu 377
626 GATACAGCCCGAGCTGGACAGATCGGCAATGCGCGGAGCGCTTCAAC 675
378 ThrThrProGluSerAlaProGluSerValProGluSerAla 394
676 GGCACTGCAGATATCGTCAAAACATCATCGCGCGGCGGAGGAAATGT 725
394 aProGluThrGluThrGluThrProThrAlaHisLeuThrThrThr 411
726 CGGCGCAGCGATGCGCGTACGGGTATAAGCGAAGGCTCAACATGCTG 775
411 hrAlaGlnThr.....ThrThrValIleThrValThr 421
776 TCATGC..... 781
422 SerCysSerAsnAlaCysSerLysThrGluValThrThrGlyValVa 438
782 .....ACGCTTGGTCTGCTTCCACCGAAACAA 812
438 lValValThrSerGluAspThrIleThrThrPheCysProLeuThr. 454
813 GATGGCGCGCATCAACGATTGGCAGATATGGCGCAACTCAAAGACTATG 862
455 .....GluThrThr 457
863 CGCAGCAGCCATCCGCGATTGGGAGTCCAAACCCCAATGCCGACAA 912
458 ProValProSerSerValAspSerThrSerValThrSerAlaProGlu 474
913 GCATAGAGCGCTCAGCAATATCTTTATGGCAGCA..... 949
474 rThrProGluSerThrAlaProGluSerSerAlaProGluSerAlaP 491
950 .....TCCCATCAAAAGGATTGGAGCTGTCGGGGGAAATACG 988
491 roGluSerSerAlaProValThrGluThrProThrGlyProValSerThr 507
989 GCTTGGCGGCGATCAGGCAATCTGTGCA.....AGCGGTGCG 1026
508 ValThrGluGlnSerLysThrIleValThrIleThrSerCysSerAsn 524
1027 CAGATGGCGCGATCGCATTCGCGAAGGA..... 1057
524 nAlaCysSerGluSerLysValThrThrGlyValValValThrSerG 541
1058 ..AATCCGCGCTCAGCGACAATTTTCCGATGCGGCAT...ACGCCAAAT 1102
541 luAspThrValThrThrPheCysProLeuThrGluThrProAla 557
1103 ACCGTGCCCTTACCATTCGGAATATCCGTTCAAACTTGGAGGAGGT 1152
558 ThrGlu.....SerAlaProGluSerSerAlaProAlaThrGluSerVa 572
1153 TAGCGAAGAAACATCACTCTCTCAACCG 1183
572 lProAlaThrGluSerAlaProValAlaPro 582
seq_name: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:US-08-616-844-40
seq_documentation_block:
; Sequence 40, Application US/08616844
; Patent No. 5849578
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A.
; TITLE OF INVENTION: COMPOSITION AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
```

```
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US 08/616,844
; FILING DATE: 15-MAR-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654
; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A.
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-053
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-616-844-40
alignment_scores:
Quality: 113.00 Length: 383
Ratio: 0.628 Gaps: 17
Percent Similarity: 46.997 Percent Identity: 21.149
alignment_block:
US-09-303-518D-463 x US-08-616-844-40 ..
Align seg 1/1 to: US-08-616-844-40 from: 1 to: 1481
110 TCGACCGCTCAGCATTCGAACCCGACGGAATACCACTATTTCGGCAGC 159
686 SerTyrIleYsIleSerAsn...SerSerHisSerGluTyrSerSerPh 701
160 AGGGGGGAGCTTGCNAGCGCAACGCGCATATCGGATTGGGAACATACA 209
701 eSerHisAlaGlnThrGluArgSerAsnIleSerSerTyrAsp..... 715
210 AAGCCATCAGTTGGCCACCTGATGATTCACAGCGCGCGTTGAAGGAA 259
716 .....GlyGlu 717
260 ATATCGGCTACATGTCGCTTTTCGATCAGCGGCACAAATTCATTTCG 309
718 TyrAlaGlnProSerThrGluSerProValLeuHisThrSerAsnLeu.. 733
310 CCCTTCGACACCACTGCTCAGATTCGATTCGACGAAGCCGCTAGTCC 359
734 .ProSerTyrThr...ProThrIleAsnMetProAsnThrSerValVal 749
360 CGTTGACGATTACGCTTTTACCGCATCCATTGGGACGATGACAAACC 409
749 euAspThrAspAlaGluPheValSerSerSerSerSerSerSerSer 765
```

```
410 ATCCCGCCGACGGCTATGACGGCCACAGCGCGGCTATCCGCTCC 459
      : : : : :
766 SerSerSerSerSerSerGlyPro.....ProLeuPr 777
      : : : : :
460 AAGGCGCGAGGATATATACACTAGCAGATAAAGGGCTGCCAAA 509
      : : : : :
777 o.....LeuProSerV 781
      : : : : :
510 TATCC.....GCCTCAACCTGACCGACACCGCAGCA 541
      : : : : :
781 alSerGlnSerHisHisLeuPheSerSerIleLeuProSerThrArgAla 797
      : : : : :
542 CCGGACAAACGGCTTGGCGAGCGTTTCCACAATGCCGCGCTATGCTGAGC 591
      : : : : :
798 SerValHisLeuLeuLysSerThrSer..... 806
      : : : : :
592 CAAGGAGTAGCGACGATCAACACCGCCGATACAGCCCGAGCT 641
      : : : : :
807 .....AspAlaSerThrProTrpSerSerProSerP 818
      : : : : :
642 GGACAGATCGGCAATGCCGCCGACCTCAACGGCACTGCAGATACG 691
      : : : : :
818 roLeuProValSerLeuThrThrSerThrSerAlaProLeuSerValSer 834
      : : : : :
692 TCAAAACATCATCGCGCGGAGGAGAAATGTGCGCGCAGCGCATGCC 741
      : : : : :
835 GlnThrThr.....LeuProGlnSerSerSerThrProValLeuPr 848
      : : : : :
742 GTCGAGGTATAGCAAGGCTCAACATGCTGTCATCGCGGCTGGG 791
      : : : : :
848 o.....ArgAlaArgGlu..... 852
      : : : : :
792 TCTGCTTTCACCGAAACAGATGCGCGCATCAACAGTTGGCAGATA 841
      : : : : :
853 .....ThrProValThrSerPheGlnThrSerThrMetThrSerPhe 866
      : : : : :
842 TGGCGC.....AATCAAGACTAT 861
      : : : : :
867 MetThrMetLeuHisSerSerGlnThrAlaAspLeuLysSerGlnSerTh 883
      : : : : :
862 GCGCAGCAGCCATCGCGATGGGAGTCCAAACCCCA..... 901
      : : : : :
883 rProHisGlnGluValIleThrGluSerLysSerProSerLeuValS 900
      : : : : :
902 ..ATGCGGCACAGGATAGAGCCGCTCAGCAATATCTTTATGGCAGCA 949
      : : : : :
900 erLeuProThrGluSerThrLys.....AlaValThrThrAsnSerPro 914
      : : : : :
950 TCCCCATCAAGGATGGAGCTGTCGCGGGAATAATACGCTTGGCGGC 999
      : : : : :
915 LeuProProSerLeuThrGluSerSerThrGluGlnThrLeuProAlaTh 931
      : : : : :
1000 ATCAGCGCATCTCTCAAGCGGTGCGAGATGGGCGCATCGCATGCC 1049
      : : : : :
931 rSerThrAsnLeuAlaGlnMetSerProThrPhe..... 942
      : : : : :
1050 GAAAGGAAATCGCGCTCAGCGACAAATTTG.....CCGATG 1087
      : : : : :
943 .....ThrThrThrIleLeuLysThrSerGlnProLeu 953
      : : : : :
1088 CGGCATACGCCA.....AATACCGTCCCTTACCATTCCCGA 1125
      : : : : :
954 MetThrThrProGlyThrLeuSerSerThrAlaSerLeuValThrGlyPr 970
      : : : : :
1126 AATATCGTTCAACCTGGCAGCGTTTACGCCAAGAAACATCACCT 1174
      : : : : :
970 olleAlaValGlnThrAlaGlyLysGlnLeuSerLeuThrHisPro 986
      : : : : :
seq_name: /cgn2_6/ptodata/1/1aa/5B_COMB.pep:US-08-599-654-40
```

seq\_documentation\_block;

; Sequence 40, Application US/08599654

```
; Patent No. 582925
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,654
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 40:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1481 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; US-08-599-654-40
```

## alignment\_scores:

|                     |        |                   |        |
|---------------------|--------|-------------------|--------|
| Quality:            | 113.00 | Length:           | 383    |
| Ratio:              | 0.628  | Gaps:             | 17     |
| Percent Similarity: | 46.997 | Percent Identity: | 21.149 |

## alignment\_block:

US-09-303-518D-463 x US-08-599-654-40 ..

Align seg 1/1 to: US-08-599-654-40 from: 1 to: 1481

```
110 TCGACCGCTCAGCATTTGCAAGCCCGAGGAAATACCACTATTTCGGCAGC 159
      : : : : :
686 SerTyrIleLysIleSerAsn...SerSerHisSerGluTyrSerSerPh 701
      : : : : :
160 AGGGGGAGCTTGCNAGCGCAACGGCCATATCGGATGGGAAACATACA 209
      : : : : :
701 eSerHisAlaGlnThrGluArgSerAsnIleSerSerTyrAsp..... 715
      : : : : :
210 AAGCCATCAGTTGGCCACCTGATGATTCACAGCGCGCTTCAAGGAA 259
      : : : : :
716 .....GlyGlu 717
      : : : : :
260 ATATCGGCTACATTGCTCGCTTTTCCGATCAGCGGCACAAATTCCTCG 309
      : : : : :
718 TyrAlaGlnProSerThrGluSerProValLeuHisThrSerAsnLeu... 733
      : : : : :
310 CCCTTCGACAAACCATGCCTCACATTCGATTCGACGACCGGCTAGTCC 359
```



Mon Jul 1 09:25:46 2002

```
1088 CGGCATACGCCA.....AATACCGTCCCTTACCATTCCTCCGA 1125
|||||
954 MetThrThrProGlyThrLeuSerSerThrAlaSerLeuValThrGlyPr 970
|||||
1126 AATATCCGTTCAAACTTGGACGACGTTACGGCAAGAAACATCACCT 1174
|||||
970 oileAlaValGlnThrThrAlaGlyLysGlnLeuSerLeuThrHisPro 986

260 ATATCGGCTACATGTCCGCTTTCCGATCAGCGGCACAAATTCATTCG 309
::: |||:::
718 TyrAlaGlnProSerThrGluSerProValLeuHisThrSerAsnLeu.. 733
|||||
310 CCCTTCGACAAACCATCCCTCACATTCGGATTCTGACGAAGCCGGTAGTCC 359
|||||
734 .ProSerTyrThr...ProThrIleAsnMetProAsnThrSerValValL 749
|||||
360 CGTTGACGGATTACGCTTTACCGCATTCATTCGGACGGATACGAACACC 409
|||||
749 euAspThrAspAlaGluPheValSerSerSerSerSerSerSerSer 765
|||||
410 ATCCCGCCGACGGCTATGACGGCCACAGGGCGGGCTATCCCGCTCC 459
::: |||
766 SerSerSerSerSerGlyPro.....ProLeuPr 777
|||||
460 AAAGGCCGAGGATATATACACCTACGACATAAAAGCGTTGCCCAAAA 509
|||||
777 o.....LeuProSerV 781
|
510 TATCC.....GCCTCAACCTGACCGACAAACCGCAGCA 541
::: |||
781 alSerGlnSerHisHisLeuPheSerSerIleLeuProSerThrArgAla 797
|||||
542 CCGGACAAACGGCTTCCGCGCCGTTTCCACAATGCCGGCGCTATGCTGACG 591
::: |||
798 SerValHisLeuLeuLysSerThrSer..... 806
592 CAAGGAGTAGCGGACGGATTCAACAGCGCCACCGCATACAGCCCGAGCT 641
|||||
807 .....AspAlaSerThrProTrpSerSerProSerP 818
642 GGACAGATCGGGCAATCCGCCGAGCCTTCAACGGCACTGCGAGATACG 691
::: |||
818 roLeuProValSerLeuThrThrSerThrSerAlaProLeuSerValSer 834
692 TCAAAACATCATCGCGCGCGAGAGAAATGTCGGCGCAGCGCATGCC 741
::: |||
835 GlnThrThr.....LeuProGlnSerSerThrProValLeuPr 848
742 GTGCGAGGTATPAGCGAGGCTCAACATTCCTGCTCATGACCGGCTTGGG 791
::: |||
848 o.....ArgAlaArgGlu..... 852
792 TCTGCTTTCCACCGAAACAGATGGCGCCATCAACGATTGGCAGATA 841
||| |||
853 .....ThrProValThrSerPheGlnThrSerThrMetThrSerPhe 866
842 TGGCGC.....AACTCAAAGACTAT 861
::: |||
867 MetThrMetLeuHisSerSerGlnThrAlaAspLeuLysSerGlnSerTh 883
862 GCCCGACGAGCATCCGGGATTGGGAGTCCAAACCCCA..... 901
|||||
883 rProHisGlnGluLysValIleThrGluSerLysSerProSerLeuVals 900
::: |||
902 ..ATCCGCACAAAGCATAGAGCGTCCAGCAATATCTTTATGGCAGCCA 949
::: |||
900 erLeuProThrGluSerThrLys.....AlaValThrAsnSerPro 914
950 TCCCATCAAAAGGATTCGAGCTGTCGGGAAATACGGCTTGGGGCGC 999
||| ::
915 LeuProProSerLeuThrGluSerSerThrGluGlnThrLeuProAlaTh 931
1000 ATCAGCGCACATCCTGCTCAAGCGGTCGAGATGGCGCGGATCGCATTCGC 1049
|||
931 rSerThrAsnLeuAlaGlnMetSerProThrPhe..... 942
1050 GAAAGGAAATCCGCGTCAGCGACAATTTTG.....CCGATG 1087
::: |||
943 .....ThrThrThrIleLeuLysThrSerGlnProLeu 953
```